

**Explaining Human Diversity:
The Need to Balance Fit and Complexity**

Armin Schulz

Department of Philosophy

University of Kansas

Lawrence, KS 66045

awschulz@ku.edu

<http://people.ku.edu/~a382s825/>

**Explaining Human Diversity:
The Need to Balance Fit and Complexity**

Abstract

While the existence of human cognitive and behavioral diversity is now widely recognized, it is not yet well established how to explain this diversity. In particular, it is still unclear how to determine whether any given instance of human cognitive and behavioral diversity is due to a common psychology that is merely “triggered” differently in different bio-cultural environments, or whether it is due to deeply and fundamentally different psychologies. This paper suggests that, to answer this question, we need to employ subtle theoretical considerations of theory choice—especially the consideration of the complexity-weighted differential predictive successes of the two accounts. To make this clearer, the paper develops a novel analysis of the observed differences in human sharing dispositions.

Keywords: human diversity; explanation; prediction; accommodation; simplicity; fit; cultural variation; WEIRD; sharing norms; evoked differences

**Explaining Human Diversity:
The Need to Balance Fit and Complexity**

I. Introduction

It is now widely recognized that human cognitive and behavioral diversity is pervasive.¹ From sharing dispositions to perceptual processing, from trading to morality, it has been found that different groups of humans systematically think and act differently (Henrich, 2000; Henrich et al., 2010; Henrich, 2020).² However, what is still not clear is exactly how to explain what gives rise to these instances of human diversity. While it is often acknowledged that that human diversity could be of two different types—human cognitive and behavioral differences could be due to a common psychology that is merely “triggered” differently in differences bio-cultural environments, or they could be the result of genuinely different psychologies—exactly how to determine which of these explanations is most plausible in which cases has not been well established. This paper looks to make progress in answering this question.

To do this, the paper focuses on one of the earliest and still most theoretically influential cases of human cognitive and behavioral variation: namely, cultural differences in sharing dispositions. Here, the paper shows that the argument for the common view that there is little psychic unity in sharing dispositions turns out to be significantly more complex than is typically supposed. In particular, the paper shows that this argument needs to employ subtle theoretical considerations of theory choice—especially the consideration of the complexity-weighted differential predictive successes of the two accounts. While the argument may end up being plausible still, this is far

¹ Though see Knobe (2019), but see also below in section V.

² The focus in what follows is on cross-cultural differences in cognitive and behavioral traits. However, the arguments can be easily extended to differences across other population groups (e.g. different genders).

from a forgone conclusion. Importantly, also, the paper shows that these lessons from the investigation of differences in sharing norms generalize to other cases of human cognitive and behavioral variation.

The paper is structured as follows. In section II, I defend and refine an old but often overlooked distinction between evoked and fundamental differences in cognition and behavior. In section III, I lay out the classic findings concerning differences in sharing dispositions, apply the evoked / fundamental distinction to this case, and show that the resolution of this case requires consideration of complex principles of theory choice. In section IV, I generalize the lessons derived in the previous section to other instances of human cognitive and behavioral diversity. I conclude in section V.

II. Two Types of Human Diversity: Evoked and Fundamental

It has long been argued that not all cases of human diversity are created equal (see e.g. Tooby & Cosmides, 1992; Boyd & Richerson, 2005; Hittinger & Carroll, 2007; Penke, 2010; Barrett, 2005; Machery, 2010; Gangestad et al., 2006). People could act and think differently simply because they have different psychologies—a *fundamental* difference—or because environmental differences, in combination with a shared underlying psychology, lead them to do so—an *evoked* difference. Drawing this kind of distinction is thought to be useful for two reasons.³

³ This distinction between “fundamental” and “evoked” difference is a relatively straightforward explication of the more traditional distinction between “transmitted” from “evoked” differences. A reason for preferring to phrase the issues in terms of “fundamental” vs. “evoked” differences, rather than “transmitted” vs. “evoked” differences, is that—as will also be made clearer momentarily—non-evoked differences need not be transmitted from a prior generation or be transmitted to a succeeding generation. They could just be the result of individuals acquiring, e.g. through individual learning, different psychological mechanisms. (At any rate, instead of “fundamental,” I also frequently use the slightly more cumbersome term “non-evoked.”) Relatedly, this terminology is not meant to suggest that a fundamental difference speaks to some sort of “essential” differences among people; rather, it is just the result of different psychological mechanisms—which themselves may have been acquired by learning or even coincidentally. While fundamental differences will be deeper than evoked differences, they may thus still be quite shallow.

First and most importantly, it enables us to get clearer on the extent to which humans are psychologically (and otherwise) unified. If many or most differences turn out to be evoked, then humans might still share a deep psychic unity. By contrast, if many or most differences turn out to be fundamental, then humans have to be seen to contain much variation even at the level of the psychological mechanisms underlying their thoughts and actions. (Note that implicit in this point is the fact that the question of the psychic unity of humans turns out to be a matter of degree, and not an either / or question.)⁴

Second, drawing the evoked / fundamental distinction is taken to be useful, as it allows us to make better predictions about how humans will think and act in a variety of situations—and thus allows us to better alter these thoughts and actions (if we deem it important to do so). So, if a difference is fundamental, then bridging it requires more than just providing people with similar experiences; rather, their psychological mechanisms need to be altered to be more in line with each other. If a difference is evoked, by contrast, it may (at least sometimes) be possible to bridge it merely by providing people with similar experiences (see e.g. Abarbanell & Hauser, 2010).⁵ Given that these are very different kinds of interventions, understanding the nature of a given human difference is thus useful for making sense of how people will react to being put in various kinds of novel circumstances.

⁴ Here and in what follows, it is presumed that cognitive and behavioral differences either are evoked or not. However, it is possible to extend this framework and see differences *as more or less* evoked (e.g. by considering how different the relevant mechanisms are). Doing this is not so relevant for present purposes, though; see also Heyes (2018). Note further that the question of the exact number of cognitive and behavioral differences that are evoked or fundamental is not central here, and will not be further pursued in this paper; what matters here is just how we can determine which differences are evoked or fundamental—not how many.

⁵ Note, though, that matters here are more complex due to the fact that providing people with similar experiences may alter their psychologies, and that providing people with similar experiences need not mean that their psychological mechanisms are triggered in the same ways, as these experiences may be differentially embedded in prior experiences. Still, the general point here stands: knowing about the nature of a cognitive and behavioral difference can help us bridge that difference.

However, it also turns out that, while seemingly straightforward, the distinction between fundamental and evoked differences needs to be further refined before it can be applied to any given instance of human diversity. Most importantly, it is necessary to determine the bedrock level of analysis, the origin of which does not need to be further explained. Such a bedrock is needed so as to form a standard with which instances of human diversity can be classified as fundamental or evoked. Without such a bedrock, the distinction between evoked and fundamental differences cannot be drawn (see also Henrich et al., 2010, pp. 59-60).

To see this, consider genetic clones of the corn plant that grow to different heights depending on whether they are in a nutrient rich or a nutrient poor soil. Are these differences in height fundamental or evoked? It may seem obvious that they are evoked: underlying these differences are not differences in the genetically-driven mechanisms of plant growth, but differences in the inputs to these mechanisms only. However, this depends on seeing genetic differences as the bedrock that grounds the analysis here. This is not the only option: in particular, if a strongly extended view of organismal development is adopted (Griffiths & Gray, 1994; Griffiths & Stotz, 2018), the difference between corn-plant-in-nutrient-rich-soil and corn-plant-in-nutrient-poor-soil may well be classified as fundamental, due to it marking two quite different *developmental systems*. Similarly, in the other direction, the fact that genetically distinct versions of the same corn plant grow to different heights in the same type of (nutrient-rich, say) soil need not be taken to imply that they are fundamentally different. If the bedrock level of analysis is seen to be the wider biological structures underwriting development in corn plants—miosis, mitosis, photosynthesis, gene expression, etc.—the genetic differences could be seen as merely (more or less accidentally determined) inputs into the fundamentally same biological mechanisms underlying corn plant development.

The general lesson behind these points is that it is necessary to determine an appropriate ground of the analysis with which differences can be classified as fundamental or evoked. What could such a ground be, though? What determines the appropriate level of analysis that is to be used as a standard with which differences are classified as fundamental or evoked? Does this require a defense of an objective ontology of psychology or biology?

Fortunately, it is not in fact necessary to engage in deep metaphysical speculation to determine the bedrock level of analysis for the evoked / fundamental distinction. All that is needed is to determine what would be most scientifically useful for the investigation of the specific kind of (human) diversity in question. That is, the resolution of the question of the appropriate grounding for the evoked / fundamental distinction is the same as that of other methodological questions in science, and there is not one answer that needs to be appropriate for all cases. Rather, the right way to draw the evoked / fundamental distinction depends on the goals of the inquiry (which may be different in purely biological contexts like that of the corn plant from what is true in purely psychological ones—among others). In fact, it may even be different for different kinds of applications in biology or psychology (psychophysics vs. social psychology, say). This kind of context- and theory-dependence is a familiar feature of science, and not something that makes the evoked / fundamental distinction problematic (Massimi, 2018; Mitchell, 2003).⁶

In the present context—i.e. when it comes to the explanation of diversity in cognitive and behavioral traits—it is plausible to see the appropriate level of analysis as consisting in the

⁶ For example, when trying to make sense of the reasons why hawks switch from hunting mice in one patch of their territory to hunting them in another, we may choose to employ a model that assumes mice are independently distributed across patches in the local area. When modeling the reasons why hawks switch prey from mice to birds as the seasons change, we may assume that the distribution of mice in the local area is auto-correlated. This is not contradictory, as these models have different goals (Potochnik, 2010; Parker, 2020).

existence of differences in the *psychological mechanisms* giving rise to the relevant mental representations or behaviors (see also Machery, 2010).⁷ That is, in what follows, the (supposed) fact that humans think and act differently because their thoughts and behaviors stem from different psychological mechanisms will be taken to mark a fundamental difference. By contrast, the (supposed) fact that they think and act differently because the same psychological mechanisms operate over different inputs will be taken to mark an evoked difference.

The reason why this is the best way of drawing the evoked / fundamental distinction in this context is that uncovering the mental mechanisms that give rise to a particular type of behavior is precisely the goal of psychological inquiry (Henrich et al., 2010; Lieder & Griffiths, forthcoming; Allen, 2014). Post behaviorism, what cognitive, developmental, and social psychology tries to establish is which mental representations humans use to navigate the world, and exactly how they do so. Of course, there are many questions that can and should be asked about exactly what these mechanisms are, where these mechanisms come from, and how they are situated in the machinery of human biology. However, for present purposes, these questions can be left open. What we are trying to do here is to classify *psychological* differences; we should therefore look towards psychology to ground this classification.

A final, related point that needs to be made here concerns learning. The fact that a behavioral or psychological difference is due to learning, by itself, does not make it either evoked or fundamental. Rather, the question is what it is that is being learned. If what is learned is a psychological mechanism (such as a decision rule), then, if different humans learn different such mechanisms, the resultant differences in thought and action will be fundamental. While it is true

⁷ The exception will be behavioral differences that are highly stereotyped or which are highly reflexive, and where the psychological processing may thus be minimal (see Schulz, 2018b). In that case, the level of analysis can be lower. However, this will not be central in what follows, and does not affect the substance of any of the conclusions reached in this paper.

that learning is, itself, a psychological mechanism, it is still the case that the results of this learning can be different psychological mechanisms—and thus that differences that are the result of the reliance on these *latter* mechanisms are fundamental. The fact that some psychological mechanisms are themselves the ontogenetic product of another psychological mechanism does not make them identical mechanisms in their own right. By contrast, if what is learned is merely the implementation or inputs into a psychological mechanism (such as the values of the variables over which a decision rule quantifies) then the resultant differences in thought and action will be evoked. Hence, the fact that a human difference is “due to learning” does not settle the question of its fundamentality.

These points become clearer when they are applied to a specific case, as does the methodology of the resolution of the question of whether a given instance of human diversity is fundamental or evoked. The next section therefore turns to the consideration of the observed differences in sharing dispositions across cultures.

III. Human Differences in Sharing Dispositions: Evoked or Fundamental?

One of the oldest findings of behavioral and psychological diversity concerns sharing dispositions. Initially, many researchers expected that all humans would be disposed to share resources in roughly comparable ways. Henrich (2000, p. 973) describes these early approaches as follows:

“Like most efforts to model human behavior in economics, [the initial] approaches, implicitly or explicitly, make certain universalist or panhuman assumptions about the nature of human economic reasoning. That is, they assume that humans everywhere

deploy the same cognitive machinery for making economic decisions and, consequently, will respond similarly when faced with comparable economic circumstances.”

However, these assumptions turned out to be mistaken: in fact, members of different cultures appear to share resources with others in their culture in drastically different ways (Henrich et al., 2005; Henrich, 2020; Henrich et al., 2010). The initial empirical findings have since been much replicated, refined, and expanded (Henrich et al., 2010; Henrich, 2020), but the overarching lessons has remained unchanged: there is massive human diversity in sharing dispositions.

There is no question that this finding is of great scientific and political importance. So, while economists do not per se assume that people are egoistic in resources—what goes into people’s utility functions is typically left open (Hausman, 2012)—it is at least a working implicit assumption of much work in economics that people prefer more of a good to less and that they are not greatly disposed towards reducing their own consumption to increase that of others (Mas-Colell et al., 1996; Hausman, 1992). Given this, the fact that there is massive diversity in sharing dispositions is problematic on several different levels. It is not just that humans are much more inclined to share than often assumed. It is that their disposition to share is highly variable, with some cultures sharing a lot, some sharing little, and many at various places in between.⁸

Most of the empirical findings concerning variation in sharing dispositions focus on the ultimatum game or various related economic games (perhaps slightly adapted to ensure better fit to the practices of the culture in question). Participants are asked how much of windfall gain they

⁸ The observed variety in sharing dispositions also has major practical implications. Given that different groups share with each other in different ways, providing aid to others has to be done in a way that respects these differences, especially when it comes to cross-cultural aid.

are willing to offer to another participant, who can then either accept or reject this offer. Rejected offers imply no gains for any participants, and all offers have to be positive.

One of the most striking findings here is that, in some cultures, minimal offers are frequently made and accepted, whereas in other cultures even hyper-fair offers (i.e. offers of more than half of the available resources) are rejected (Henrich, 2000; Henrich et al., 2005; Henrich et al., 2010; Henrich et al., 2001; Bone et al., 2016; Blake et al., 2015). For example, the median offer among the Hadza is about 26% of the stake, but even lower offers are frequently accepted—whereas the median offer among the Sursurunga is over 50% of the stake, and lower offers are frequently rejected. In between these two extremes are many different median offer / acceptance rate combinations (see table 1):

Culture	Mean Offer (approx. %)	Income Maximizing Offer (approx. %)
Maragoli	25	40
Hadza	26	10
Tsimane	27	10
Samburu	25	10
Shuar	36	10
Isanga	38	10
Gusli	40	40
Yasawa	40	10
Nganasan	43	10
Dolgan	44	10
Au	44	20
Accra	44	20
Sanquianga	47	10
US	47	50
Sursurunga	52	50

[Table 1 (from Henrich, 2010, p. 6, figure 3(b) and (c))]

In short: from a cross-cultural point of view, we find that, in some human groups, people share a lot with each other, and expect others to share a lot with them, whereas in other human groups, people neither share nor expect others to share much.

How can we make sense of these differences in human sharing dispositions? In particular, do these differences imply that humans in different cultures really do *think* very differently about sharing? Or is it the case that humans, despite this variability, actually share in line with the same psychological mechanism, but that this shared mechanism happens to lead to different outcomes in different cultures (for reasons that need to be made clearer still)?

1. Human Diversity in Sharing Dispositions: The Two Accounts

To answer these questions, it is best to begin by getting clearer on what the fundamental and evoked accounts come down to here. Both of these accounts agree that something about the cultural environment of the different human groups in question is at the root of the differences; they just differ over what this is (see e.g. Henrich et al., 2005, pp. 841-846).

The evoked account suggests that underlying the diversity in the observed sharing dispositions is a fundamental universality. All cultures rely on the same psychological mechanisms as far as sharing is concerned; it is just that these mechanisms contain variables or parameters that (a) have different values in different cultural settings, and / or (b) which are “instrumented”—estimated using available data—differently in different cultural settings. As a result of (a) and (b), there will be different outcomes as far as the resultant sharing behavior is concerned—despite the psychological mechanisms underlying the sharing dispositions themselves being the same (see e.g. Kenrick & Sundie, 2005; Machery, 2010).

There are many forms such an evoked account could take, but the most well developed version—and the one that will be the focus of the discussion here—starts from one of the core principles of evolutionary biology: kin selection (Kenrick & Sundie, 2005; Cronk & Gerkey, 2007; Gowdy et al., 2013; Henrich, 2020). (Note also that the spirit of the conclusions reached in this paper will be maintained if another version of the evoked account is adopted.) Kin selection has come to be recognized as one of the major drivers of evolution (Gardner et al., 2011; Griffin & West, 2002; West et al., 2011; West et al., 2007, 2008; Sober & Wilson, 1998; Okasha, 2006; Birch & Okasha, 2014; Bowles & Gintis, 2011). At its heart, kin selection theory is based on the fact that giving away resources can be highly adaptively advantageous—but only in the right circumstances. These circumstances center on giving to *kin*: organisms that are genetically similar to the focal organism (Schulz, 2018a; Gardner et al., 2011; Frank, 1998). Given the wide applicability of kin selection across the biological world, it stands to reason that humans, too, are more inclined to share with kin than with non-kin (Kenrick & Sundie, 2005; Cronk & Gerkey, 2007; Gowdy et al., 2013; Henrich, 2020).

However, it is also highly likely that, in different cultures, (a) people are differentially likely to interact with kin, and (b) estimating who is kin and to what extent is differentially difficult (Cronk & Gerkey, 2007; Markman et al., 2005). For example, if all the children in the community are raised together and separately from the parents (as in a Kibbutz), or if the culture is polyandrous and children grow up with their mothers, then determining who is related to who and to what extent can be complex and needs to rely on more or less reliable cues. People in some cultures may come to learn, more or less accurately, that everyone they are interacting with very closely related to them, whereas others learn the opposite. Alternatively, people in some cultures may come to learn, more or less accurately, that they are closely related to people who

have tattoos of a certain shape, whereas people in other cultures may learn this about people who have a certain kind of accent when speaking.

The upshot of this is that there can be major differences in the *apparent* sharing dispositions of people from different cultures, even if, *fundamentally*, their sharing psychologies are identical (Hintze & Hertwig, 2016; Markman et al., 2005). In particular, it may be that all humans share by relying on a psychological mechanism of this sort:

Sharing Mechanism: “Provide person y with resources yielding b benefits to them, at a cost of c resources to me, if $rb > c$, where r is the coefficient of relatedness between me and y .”

However, it is also the case that humans need to learn to use the available cultural clues to estimate r . In turn, this can lead to different sharing outcomes, both because people might actually be differentially related to each other in different cultures, and because the estimates of their relatedness may be differentially accurate (Kenrick et al., 2008; Kenrick & Sundieb, 2005; see also Cosmides & Tooby, 1992). That is to say, the account allows different cultures C_1 , C_2 , and so on to have different instrumentations of r (r_1 , r_2 , ...). These different instrumentations, though, are constrained by the fact that they aim to accurately estimate the same (biologically important) variable.⁹

⁹ Another example of this sort of case has been provided by Henrich et al. (2005, p. 811): they suggest that the fact that even hyper-fair offers in the ultimatum game are frequently rejected among the Au and Gnaou is due to it being the case that, in this culture, the acceptance of a gift is taken to imply an obligation to repay this gift at a later date. If so, though, then this kind of rejection of hyper-fair offers should not be seen to display a different attitude towards resource division—it is just an aspect of the fact that, in this culture, gift giving is a much more dynamically extended affair that includes repayment of the gift later on. If this point is taken into account, the differences in the sharing dispositions between this culture and others might well disappear: *holding the value of a gift fixed* (which may include considering any obligations to repay the gift later), people from different cultures may display the same sharing dispositions (Kenrick & Sundieb, 2005).

By contrast, the fundamental account of cultural differences in sharing dispositions suggests that the psychological mechanisms underlying these sharing dispositions, even at a fundamental level, are learned from the cultural environment, and thus differ from each other (Henrich et al., 2005, pp. 812-814, 842-846; Henrich, 2000, p. 973). The fundamental account may agree with the evoked account that there are some non-cultural biological constraints on the nature of the psychological mechanisms that underpin the observed sharing dispositions—especially those deriving from kin selection. However, the fundamental account sees the major proportion of the variance in the transmission and adoption of these mechanisms as influenced by factors unique to specific cultures: these cultural factors are so strong as to outweigh most of the non-cultural biological constraints on sharing dispositions.

This account may appeal to the fact that similar explanations are plausible in other cases. For example, consider politeness norms. While there may be some reason to think that there are non-cultural biological constraints on the cultural spread of politeness norms—those norms that fit better to evolved human emotional endowments (e.g. in terms of disgust) are more likely to persist and spread (Nichols, 2004; Boyd & Richerson, 2005; Henrich, 2015; Henrich & McElreath, 2007)—much of this spread is due to cultural learning. Our evolved human emotional endowments may constrain this learning, but they do not constrain it so much that only a single politeness norms remains.

The same idea can be applied here. While there may well be some non-cultural biological constraints on the spread of the psychological mechanisms underlying sharing, much of this spread is to be explained directly by appeal to the cultural learning of different such mechanisms (Boyd & Richerson, 2005; Henrich, 2015; Henrich & McElreath, 2007; Bowles & Gintis, 2011).

There is a myriad of factors that could lead different cultures to learn different mechanisms for sharing. Such differences could arise from accidental differences in the models that are being copied (e.g. whether an individual with a more generous sharing psychology happens to be identified as a good model), differences in the learning mechanisms involved (e.g. whether, in the culture in question, it is common to learn from few models or from a majority of models), as well as correlations with other features of the relevant culture (how much use it makes of market interactions, how stratified it is, or how warlike it is), each of which further rests on other gene-cultural factors (Henrich, 2015; Henrich & McElreath, 2007; Boyd & Richerson, 2005; Henrich et al., 2005).

The upshot of this is that, on the fundamental account, differences in the sharing behavior of people from different cultures are largely due to the fact that people simply have different psychologies as far as sharing is concerned. There may be some constraints on the kinds of sharing dispositions that can spread in different cultures (such as ones based on kin selection), but, by and large, the psychological mechanisms underlying sharing dispositions are themselves culturally acquired:

Sharing Mechanism C₁: “Provide person y_{i1} with resources yielding b_{i1} benefits to them, at a cost of c_{i1} resources to me, by following rule R_1 .”

Sharing Mechanism C₂: “Provide person y_{i2} with resources yielding b_{i2} benefits to them, at a cost of c_{i2} resources to me, by following rule R_2 .”

Sharing Mechanism C₃: “Provide person y_{i3} with resources yielding b_{i3} benefits to them, at a cost of c_{i3} resources to me, by following rule R_3 .”

...

In this case, each culture has its own sharing psychology that is based on different rules (the R_i above), each of which specifies how many resources are to be shared with who.¹⁰ While kin selection pressures may put some constraints on what these R_i can be, they still vary widely. (Parallel remarks could be made for other versions of the evoked account.)¹¹

2. *Theory Choice in the Explanation of Differences in Sharing Dispositions*

Which of these two accounts is more plausible? Are human sharing dispositions varied only in a shallow—evoked—sense, or are they deeply—fundamentally—varied? Before it is possible to answer this question, two things need to be noted.

First and most obviously, it needs to be noted that there is a question to be asked here. The mere fact that there is human diversity in sharing dispositions does not mean that this diversity needs to be fundamental. To be sure, the fact that different groups of humans share in different ways shows that there is *some* kind of human variation; however, it does not speak to question of *what* kind of variation it is (see also Machery, 2010). This is important to note, as it is sometimes implied that documenting the existence of human diversity in sharing dispositions by itself shows that there is no such thing as human psychic unity in this regard. So, Henrich et al. (2005, p. 844) make the following remarks concerning the ultimatum game:

¹⁰ So, R_1 may specify that we share three strawberries with our nearest neighbor, and four tomatoes with our first cousins, while R_2 may specify that we share two strawberries with our nearest neighbor, and eight potatoes with our first cousins. (Obviously, this is a purely illustrative example.)

¹¹ This leaves open whether the individual rules in the different cultures—i.e. the different R_i —have individually more complexity than the common sharing mechanism (e.g. Hamilton's rule) of the evoked account. This will clearly depend on the nature of the evoked account in question. However, what matters here is just that the fundamental account specifies a different rule for each culture—with potentially a completely different set of determinant variables.

“If our fitness maximizer is the proposer and she knew that the respondent is also a fitness maximizer, she should offer the smallest positive amount possible, knowing that the fitness-maximizing respondent will accept any positive offer. This simple fitness maximizing prediction is not supported in any society. Thus, our work [i.e. the cross-cultural findings sketched above] provides an empirical challenge to evolutionary theory.”

Similarly, Henrich (2000, p. 973) say:

“[I]f the Machiguenga results [and others like them] stand the test of scrutiny and can be replicated elsewhere, then the assumption that humans share the same economic decision-making processes must be reconsidered.”

However, neither of these points in fact follow from the above data about diversity in sharing dispositions. As long as the evoked approach is understood properly, all of the documented facts of diversity in sharing dispositions are consistent with an underlying shared psychic unity in this regard. Inclusive fitness-maximizing humans may make any offer in the ultimatum game—this depends on their estimated value of relatedness to the recipient and other details of the case, such as the biological value of resource in question (Hamilton, 1964; Gardner et al., 2011; Rubin, 2018; Grafen, 2006; Sober & Wilson, 1998; Godfrey-Smith, 2008; Piccinini & Schulz, 2019).

Second and relatedly, the fact that humans are genetically very similar (Boyd & Richerson, 2005; Kenrick & Sundieb, 2005; Galanter et al., 2017) does not speak to whether their sharing dispositions are fundamental or not. On *neither* of these accounts do the observed differences in sharing dispositions have genetic underpinnings. This is important, as it has caused some

confusion in the past: for example, Gaertner et al. (2010) criticize Henrich et al. (2010) for failing to note that underlying human variability in cognitive and behavioral traits, there may be a common genotype. However, as also noted by Henrich et al. (2010, pp. 59-60), this is really not the issue here.

How can we make progress in explaining human diversity in sharing behavior, then? As it turns out, to do this, it is necessary to take into account some subtle issues of theory choice in science. The first point to note here is that the two accounts make two quite different predictions concerning people's disposition to share with kin relative to non-kin.

The evoked account predicts that people, regardless of their cultural background, are more likely to share with biological kin—holding the value and cost of the resource to be shared constant (Hamilton, 1964; Gardner et al., 2011). Of course, as just noted, this is consistent with there being some cultural differences in how people share with kin, since the “instrumentation” of kin (i.e. how biological kin are detected and conceived of by people) can differ across cultures (Cronk & Gerkey, 2007). However, the key point for present purposes is that reasons need to be provided *ex ante*—i.e. before the data are considered—for why these instrumentations lead to different outcomes. *Before looking at the data*, if there is no reason to think that people differ in their estimates of who is kin with who or of what the value the relevant resource is, the evoked account predicts that people will be disposed to share in about the same way (Kenrick & Sundieb, 2005). Of course, it is possible to posit differences in the instrumentation of kin (etc.) *ex post*, i.e. once the data have been collected. The question, though, is what the evoked account *predicts* before the data are consulted. Here, the answer is clear: if, given all that we know, cultures A and B instrument kin similarly, then, in the same circumstances, the evoked account predicts that people in cultures A and B will share in the same way. That is to say, *ex ante* and

ceteris paribus, we should expect $r_i = r_j$ in the universal *Sharing Mechanism*, for different two cultures i and j .

By contrast, it is a core feature of the fundamental account that the cultural evolution of the psychological mechanisms underlying sharing dispositions—while (potentially) checked and constrained by non-cultural biological factors—operates quite independently of these non-cultural biological features (Boyd et al., 2011; Henrich, 2015; Henrich et al., 2005; Henrich & McElreath, 2007). Of course, it is *possible* that all cultures happen to share with mechanisms that are fundamentally the same—i.e. that $R_i = R_j$ for all i and j in *Sharing Mechanism C_i* and *Sharing Mechanism C_j* above. However, as such, there is no inherent reason on the fundamental account to *predict* that this would be so *ex ante*. As above, it is of course possible to posit, *ex post*, that $R_i = R_j$ for all cultures in one's sample; it may turn out that, once all the data are collected, this is the most compelling variable assignment here. However, *ex ante*, the fundamental account *predicts* that the biologically largely unconstrained cultural learning of sharing mechanisms will show up somewhere—i.e. that $R_i \neq R_j$ for some i, j (see also Henrich et al., 2001, p. 75). This includes, inter alia, sharing with kin. Put more bluntly: the fundamental account does *not* predict, *ex ante*, that sharing with kin is culturally universal—holding fixed the way kin are instrumented in different cultures (though, as noted above, it can easily *accommodate* this fact—if it is a fact—*ex post*).

In this way, cross-cultural sharing behavior towards kin can be seen as a litmus test for which of the two theories of the nature of human sharing dispositions is more compelling. The core difference between these two accounts (whether there is cultural variability in the mechanisms underlying human sharing dispositions, or merely variability in the circumstances of the use of

these mechanisms) comes out particularly clearly when it comes to *predictions*—though not accommodations—about sharing with kin.

However, there is a second point that needs to be taken into account in the comparison between the evoked and fundamental accounts. This second point concerns the differential *complexities* of the two accounts.¹² The evoked account contains more restrictions than the fundamental account: according to it—*ex ante* and unless a reason otherwise is specified—humans in different cultures use the same r in their same sharing mechanism “Provide person y with resources yielding b benefits to them, at a cost of c resources to me, if $rb > c$, where r is the coefficient of relatedness between me and y .” By contrast—*ex ante* and unless a reason otherwise is specified—the fundamental account sees all the R_i in “Provide person y with resources yielding b benefits to them, at a cost of c resources to me, if R_i is true” to be *different*. In this way, *ex ante*, the evoked account is has fewer degrees of freedom to fit the data than the latter: it only has one variable (r_i) that has to fit all n cases, rather than having n variables (R_i , for $i = 1$ to n) for n cases.

This matters, as, from a methodological point of view of theory choice, it is not just important whether one theory fits the data better than another. Rather, it is important whether a theory fits the data *sufficiently better to make up for any increases in its complexity* (Hitchcock & Sober, 2004; Forster & Sober, 1994). After all, theories with more degrees of freedom will always be able to mimic theories with fewer degrees of freedom by setting some of their extra degrees of freedom equal to zero. Hence, if we only consider the predictive successes of rival theories, we would be led to always favor more complex theories over less complex ones.

¹² Other theoretical virtues may matter too, but these are less well understood. Hence, the focus is on simplicity / complexity here. Considering other theoretical virtues would only strengthen the arguments of this paper.

This, though, would not be compelling, as we are then in constant danger of “overfitting” the data (Hitchcock & Sober, 2004; Forster & Sober, 1994; White, 2003). Overfitting is problematic, as it results in us failing to distinguish the signal from the noise in the data. Any real world data set contains error and randomness, and the more degrees of freedom a theory has, the more likely it is that it is swayed by this kind of error and randomness, and pushed away from delineating the trend that underlies this error and randomness (Burnham & Anderson, 2002; Hitchcock & Sober, 2004; Forster & Sober, 2011, 1994; Schwarz, 1978; Rochefort-Maranda, 2016).

Now, it needs to be acknowledged that it is not always straightforward to determine exactly how complex a theory is relative to a competitor, or exactly how strongly increases in complexity should be penalized. Measuring complexity is relatively easy if one theory is a nested version of another—for example, when we compare a theory that has the form of a straight line ($y = ax + b$) with one that has the form of a parabola ($y = ax^2 + bx + c$). It is more complex, though, if the theories are not nested in each other (as is the case when it comes to sharing norms). On top of this, there is debate over exactly how to penalize increases in complexity. The Akaike Information Criterion (AIC), the Bayesian Information Criterion (BIC), and Likelihood Ratio (LHR) tests—to name just a few—all propose slightly different ways of tempering predictive success by theory complexity (Zucchini, 2000; Bretthorst, 1996; Forster & Sober, 2011; Burnham & Anderson, 2002; Schwarz, 1978; Goodman & Royall, 1988; Royall, 1997; Stone, 1974, 1977).¹³

For present purposes, though, these details of how to measure and evaluate theory complexity do not matter. What matters here is just that, on the one hand, while not always straightforward,

¹³ These different frameworks also differ in terms of their requirements. For example, AIC requires strictly nested models, whereas BIC or LHR-based methodologies allow for comparisons among non-nested models (Burnham & Anderson, 2002; Abraham & Ledolter, 2006; Smith, 1992).

the complexities of the fundamental and evoked accounts can be compared. To do this, the two accounts need to be spelled out in sufficient detail and placed in a suitable model selection framework that allows the assessment of their differential predictive—not accommodative—success, adjusted for their complexity. This is no different than in other cases of theory choice in science (Forster & Sober, 1994; Hitchcock & Sober, 2004; Sober, 1988; Rochefort-Maranda, 2016; Schulz, 2018a; Ruiz & Schulz, forthcoming). On the other hand and most importantly, this kind of comparison *needs to be made*. It cannot be ignored or skipped: without it, the assessment of whether the evoked or fundamental accounts are more plausible cannot be compellingly made. This is important, as the existing literature on the investigation of human diversity has tended to fall into exactly this trap: it never even considers the question of the complexity of the two accounts.

Moving beyond the existing literature, therefore, the following points about the relative plausibility of the evoked and fundamental accounts can be made. In the first place, there is at least some reason to think that the fundamental account has a higher predictive accuracy than the evoked account. Given that the evoked account predicts that people across cultures should be more inclined to share with biological kin (*ceteris paribus*), it should be the case that more close-knit, family-based cultures should show higher degrees of sharing in the ultimatum game. In these cultures, by definition, most sharing happens among kin, so that, unless reasons for the differential instrumentation of the universal sharing mechanism are provided—which has not been done—people should be expected to work with the heuristic that sharing even with an anonymous member of their culture involves sharing with a relatively close kin (Kenrick & Sundieb, 2005). However, this is not what we find: there is much variation in the sharing dispositions even among close-knit, family-based cultures, with some cultures showing very

little sharing behavior. For example, the mean offer among the Achuar is around 40% of the stake, while it is only about 25% among the Quichua (Henrich et al., 2005, p. 801). This thus speaks against the prediction of the evoked account.

This point is further strengthened by the fact that a number of studies have shown that kinship is less relevant to cooperative decisions generally in many cultures than cultural factors such as “lineage-membership” (in matrilineal or patrilineal societies) (Alvard, 2003; Cronk & Gerkey, 2007). For example, in a study of the ultimatum game among the Bwa Mawego (Macfarlan & Quinlan, 2008), it was found that “average relatedness to the village did not predict the size of proposals” (Macfarlan & Quinlan, 2008, p. 304), and that “[t]he matrifocal nature of village life may pattern the altruism received across the life course such that males and females develop different sets of models of fairness conditional on family characteristics” (Macfarlan & Quinlan, 2008, p. 306). Similarly, whale-hunting crews among the Lamalera—which need and tend to be highly cooperative—are formed, in the first instance, of members of the same lineage, not kin. While this latter study, and some others like it (Cronk & Gerkey, 2007), do not consider the ultimatum game directly, they at least provide reasons to think that sharing is not always kin-driven.

Now, these points need to be corroborated with a more systematic meta-review in order to conclude that the fundamental account really has the higher predictive success here. This is especially so since there are also studies that do not fit the predictions of the fundamental account as well: for example, kinship relations seem the key underlying factor in cooperative behavior among Mosuo farming households in China (see e.g. Thomas et al., 2018; see also Macfarlan & Quinlan, 2008, p. 308). Still, at least as matters stand, when it comes the prediction of differences in sharing dispositions, the fundamental account seems to come out slightly ahead.

However, it also needs to be noted that the fundamental account has a higher degree of complexity. While the evoked account may need to add variables for some divergences in how kin are instrumented in different cultures, overall (as noted earlier), it is more restricted than the fundamental account. Whether this increase in complexity outweighs the higher predictive success of the fundamental account is not yet clear, and will depend on how all the data here turn out, as well as what the best model selection framework for the case is. However, what can be concluded at this point is that the case for the view that human dispositions to share are fundamentally influenced by cultural factors is far from obvious. It may end up being the right view—but whether that is so cannot yet be confidently asserted. Moreover, if this case does end up being plausible, it will rest on subtle differences in the predictive successes of the two accounts, taking into account their complexities. It will not be an obvious question that can be settled with a quick look at the data.

While this conclusion is interesting in its own right, for present purposes, it also has value in illustrating points that generalize to the investigation of human cognitive and behavioral diversity more broadly. The next section makes this clearer.

IV. Theory Choice and the Investigation of Human Cognitive and Behavioral Diversity

There are two lessons that need to be kept in mind when investigating cases of human diversity in cognitive and behavioral traits quite generally. First and most obviously, the mere fact that there is human diversity in some trait does not speak against the psychic unity of humans. This diversity may turn out to be merely evoked, and thus, fundamentally, humans may be alike in the relevant regard. As noted earlier, while sometimes noted, this point still continues to be frequently overlooked, and thus deserves stressing: both the evoked and the fundamental

accounts are equally consistent with the fact that there is variation in the ways in which humans think or act.¹⁴

However, second, this does not mean that the question of the extent of human psychic unity is intractable. Rather, whether differences in behavior should be seen to be evidence for differences in the psychological mechanisms underlying and causing this behavior turns on subtle, but also manageable and known matters of theory choice. In particular, the choice between the evoked and the fundamental accounts is a matter of the evaluation of two competing scientific theories. In this way, it becomes possible to answer the worry of Henrich et al. (2010, p. 60), when they note (in response to Gaertner et al., 2010):

“It seems to us that Gaertner et al. are offering an unfalsifiable hypothesis. They suggest that studying diverse populations will either yield evidence of similarities because of an underlying universal psychological process, or it will yield evidence of differences, which mask the underlying universal psychological process. They do not offer any means for discerning an underlying universal process in the face of population-level variability.”

Henrich et al. (2010) may be right that Gaertner et al. (2010) do not specify “any means for discerning an underlying universal process in the face of population-level variability.” However, in general, determining whether the fundamental or the evoked account provides a more compelling explanation of a given type of observed variation in some cognitive or behavioral

¹⁴ Note also that the points made here are different from those in the debate between Knobe (2019) and Machery and Stich (forthcoming). At stake in the latter debate is the question of how much variation there is in various behavioral and cognitive traits, with the former arguing that it is less than often supposed, and the former arguing that it is more. By contrast, the point here is just that, even if it turns out that diversity in psychological and behavioral traits is pervasive, this does not mean it is also fundamental. So, independently of how the Machery & Stich and Knobe debate is being resolved, the question of the psychic unity still needs to be answered.

trait is not a mystery. Rather, it depends on an assessment of the differential predictive success of the two accounts, weighted by the complexities of the two accounts.

Now, as noted in the previous section, it is true that this kind of assessment is complex on many different levels. The comparison of the differential predictive success of the two theories need not be obvious or unequivocal, as the two theories may make only subtly different predictions (e.g. one that differ in quantity), and which it need not be easy to empirically conform. Similarly, the assessment and comparison of the complexity of the two theories is also something that need not be entirely without complexities. However, importantly, the evaluation of whether an instance of human diversity in cognitive or behavioral traits is evoked or fundamental is also far from impossible. It takes spelling out the evoked and fundamental accounts in sufficient detail so that their differential predictions can be checked and their complexity assessed, as well as conducting careful empirical evaluations of many cases of human diversity. While this can be tricky and will often not yield unequivocal answers (at least for a while)—as is true when it comes to human diversity in sharing norms—this is not different from other cases of theory evaluation in science.

This matters, as it holds some immediate lessons for the investigation of other instances of human cognitive and behavioral diversity. So, for example, in the discussion of human diversity in moral cognition, it has been pointed out that underneath this variation in moral thought and action could be a common, unified set of psychological mechanisms (Mikhail, 2011; Sripada, 2008; Abarbanell & Hauser, 2010—though these different authors have different views about what this common set of mechanisms is).¹⁵

¹⁵ Note that moral cognition may well turn out not to be one trait, but several: humans may rely on many different psychological mechanisms, each of which is tailored to a different moral domain or issue. However, this is not central here, and does not affect the conclusion reached.

To determine whether this is plausible, it is not enough to ask whether the evoked or the fundamental account can accommodate a given data point, such as the fact that many or all cultures seem to have an incest taboo, or that some, but not all cultures, allow for cannibalism. Every sufficiently well spelled out account—whether evoked or fundamental—will likely be able to accommodate all the data. What needs to be asked, is how well the accounts do at predicting the data while also considering how complex they are. This is not easy, but nor is it impossible. (Very similar remarks can be made concerning the discussion of differences in epistemic concepts: see e.g. Starmans & Friedman, 2020.)

V. Conclusion

The fact that cultural influences can have major effects on cognition and behavior is, undoubtedly, important and something that needs to be widely recognized. Too often, it is still glibly assumed that, because a given person and their inner circle think and act in a certain way, that every human does so, too. Correcting this mistaken inference is important.¹⁶ However, it is also important to recognize that the issues here are quite complex.

In particular, the fact that cultural influences can have major effects on cognition and behavior—important as this fact is—cannot automatically be taken to mean that humans are not fundamentally psychologically alike. As this paper makes clear, overall, the extent to which there is human psychic unity or not is still an open question. To answer it, it needs to be determined whether each case of human diversity is evoked or fundamental. In turn, in order to do this, the predictions of the fundamental explanation—i.e. the predictions based on assumed the observed

¹⁶ Note that this is a point that Knobe (2019) recognizes, too: his claim is precisely that it is *very surprising* that it *empirically* turns out that many people think alike in many ways. While the latter part of this claim is being denied by Machery and Stich (forthcoming), the former part clearly speaks to the fact that pervasive and obvious human psychic unity is not something that should be presumed from the armchair.

difference is fundamental—first need to be compared to the evoked explanation—i.e. the predictions based on assumed the observed difference is evoked. Second, the differential predictive success of the two accounts then needs to be weighted by their differing complexities. There is also no question that this is a difficult research project to carry out, and that it is unlikely to yield straightforward answers. However, neither is this an impossible project to engage in—as the case of differences in sharing dispositions shows. Without a doubt, it is an important project that can tell us much about the unique type of creatures that humans are.

Bibliography

- Abarbanell, L., & Hauser, M. D. (2010). Mayan morality: an exploration of permissible harms. *Cognition*, *115*(2), 207-224.
- Abraham, B., & Ledolter, J. (2006). *Introduction to Regression Modeling* (1st ed.). Independence, KY: Cengage.
- Allen, C. (2014). Models, Mechanisms, and Animal Minds. *The Southern Journal of Philosophy*, *52*, 75-97.
- Alvard, M. S. (2003). Kinship, lineage, and an evolutionary perspective on cooperative hunting groups in Indonesia. *Human Nature*, *14*(2), 129-163.
- Barrett, H. C. (2005). Adaptations to Predators and Prey. In D. M. Buss (Ed.), *The Handbook of Evolutionary Psychology* (pp. 200-223). Hoboken, NJ: Wiley.
- Birch, J., & Okasha, S. (2014). Kin Selection and Its Critics. *BioScience*, *65*(1), 22-32.
- Blake, P. R., McAuliffe, K., Corbit, J., Callaghan, T. C., Barry, O., Bowie, A., . . . Warneken, F. (2015). The ontogeny of fairness in seven societies. *Nature*, *528*, 258.
- Bone, J. E., McAuliffe, K., & Raihani, N. J. (2016). Exploring the Motivations for Punishment: Framing and Country-Level Effects. *PloS one*, *11*(8), e0159769.
- Bowles, S., & Gintis, H. (2011). *A Cooperative Species: Human Reciprocity and Its Evolution*. Princeton: Princeton University Press.
- Boyd, R., & Richerson, P. (2005). *The Origin and Evolution of Cultures*. Oxford: Oxford University Press.
- Boyd, R., Richerson, P., & 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.

- Bretthorst, G. L. (1996). An Introduction to Model Selection Using Probability Theory as Logic. In G. Heidbreder (Ed.), *Maximum Entropy and Bayesian Methods* (Vol. 62, pp. 1-42): Springer Netherlands.
- Burnham, K. P., & Anderson, D. R. (2002). *Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach* (2nd ed.). New York: Springer.
- Cosmides, L., & Tooby, J. (1992). Cognitive adaptations for social exchange. In J. Barkow, L. Cosmides & J. Tooby (Eds.), *The Adapted Mind: Evolutionary Psychology and the Generation of Culture* (pp. 163-228). Oxford: Oxford University Press.
- Cronk, L., & Gerkey, D. (2007). Kinship and Descent. In R. Dunbar & L. Barrett (Eds.), *The Oxford Handbook of Evolutionary Psychology* (pp. 463-478). Oxford: Oxford University Press.
- Forster, M., & Sober, E. (1994). How to tell when simpler, more unified, or less ad hoc theories will provide more accurate predictions. *British Journal for the Philosophy of Science*, 45(1), 1-35.
- Forster, M., & Sober, E. (2011). AIC Scores as Evidence—A Bayesian Interpretation. In M. Forster & P. S. Bandyopadhyay (Eds.), *The Philosophy of Statistics* (pp. 535-549). Dordrecht: Kluwer.
- Frank, S. A. (1998). *Foundations of Social Evolution*. Princeton: Princeton University Press.
- Gaertner, L., Sedikides, C., Cai, H., & Brown, J. D. (2010). It's not WEIRD, it's WRONG: When Researchers Overlook uNderlying Genotypes, they will not detect universal processes. *Behav Brain Sci*, 33(2/3), 33-34.

Galanter, J. M., Gignoux, C. R., Oh, S. S., Torgerson, D., Pino-Yanes, M., Thakur, N., . . .

Zaitlen, N. (2017). Differential methylation between ethnic sub-groups reflects the effect of genetic ancestry and environmental exposures. *eLife*, 6.

Gangestad, S. W., Haselton, M. G., & Buss, D. M. (2006). Evolutionary foundations of cultural variation: Evoked culture and mate preference. *Psychological Inquiry*, 17, 75-95.

Gardner, A., West, S. A., & Wild, G. (2011). The genetical theory of kin selection. *J Evol Biol*, 24(5), 1020-1043.

Godfrey-Smith, P. (2008). Varieties of Population Structure and the Levels of Selection. *British Journal for the Philosophy of Science*, 59, 25-50.

Goodman, S. N., & Royall, R. (1988). Evidence and Scientific Research. *American Journal of Public Health*, 78(12), 1568-1574.

Gowdy, J., Rosser Jr., J. B., & Roy, L. (2013). The evolution of hyperbolic discounting: Implications for truly social valuation of the future. *Journal of Economic Behavior & Organization*, 90S, S94-S104.

Grafen, A. (2006). Optimization of Inclusive Fitness. *J Theor Biol*, 238, 541-563.

Griffin, A. S., & West, S. A. (2002). Kin Selection: Fact and Fiction. *Trends in Ecology and Evolution*, 17, 15-21.

Griffiths, P., & Gray, R. D. (1994). Developmental Systems and Evolutionary Explanation. *The Journal of Philosophy*, 91(6), 277-304.

Griffiths, P., & Stotz, K. (2018). Developmental Systems Theory as a Process Theory. In D. J. Nicholson & J. Dupre (Eds.), *Everything Flows: Towards a Processual Philosophy of Biology* (pp. 225-245). Oxford: Oxford University Press.

Hamilton, W. (1964). The Genetical Theory of Social Behavior. *J Theor Biol*, 7, 1-52.

- Hausman, D. M. (1992). *The Inexact and Separate Science of Economics*. Cambridge: Cambridge University Press.
- Hausman, D. M. (2012). *Preference, Value, Choice, and Welfare*. Cambridge: Cambridge University Press.
- Henrich, J. (2000). Does Culture Matter in Economic Behavior? Ultimatum Game Bargaining Among the Machiguenga of the Peruvian Amazon. *American Economic Review*, 90(4), 973-979.
- Henrich, J. (2015). *The Secret of Our Success: How Culture Is Driving Human Evolution, Domesticating Our Species, and Making Us Smarter*. Princeton, NJ: Princeton University Press.
- Henrich, J. (2020). *The WEIRDest People in the World*. New York: Farrar, Straus and Giroux.
- Henrich, J., Boyd, R., Bowles, S., Camerer, C., Fehr, E., Gintis, H., & McElreath, R. (2001). In Search of Homo Economicus: Behavioral Experiments in 15 Small-Scale Societies. *The American Economic Review*, 91(2), 73-78.
- Henrich, J., Boyd, R., Bowles, S., Camerer, C., Fehr, E., Gintis, H., . . . Tracer, D. (2005). “Economic man” in cross-cultural perspective: Behavioral experiments in 15 small-scale societies. *Behavioral and Brain Sciences*, 28(6), 795-815.
- Henrich, J., Heine, S. J., & Norenzayan, A. (2010). The weirdest people in the world? *Behav Brain Sci*, 33(2-3), 61-83; discussion 83-135.
- Henrich, J., & McElreath, R. (2007). Dual-Inheritance Theory: The Evolution of Human Cultural Capacities and Cultural Evolution. In R. Dunbar & L. Barrett (Eds.), *The Oxford Handbook of Evolutionary Psychology* (pp. 555-570). Oxford: Oxford University Press.

- Heyes, C. M. (2018). *Cognitive Gadgets: The Cultural Evolution of Thinking*. Cambridge, MA: Harvard University Press.
- Hintze, A., & Hertwig, R. (2016). The Evolution of Generosity in the Ultimatum Game. *Scientific Reports*, 6, 34102.
- Hitchcock, C., & Sober, E. (2004). Prediction Versus Accommodation and the Risk of Overfitting. *The British Journal for the Philosophy of Science*, 55(1), 1-34.
- Hittinger, C. T., & Carroll, S. B. (2007). Gene duplication and the adaptive evolution of a classic genetic switch. *Nature*, 449(7163), 677-681.
- Kenrick, D. T., Sundie, J. M., & Kurzban, R. (2008). Cooperation and conflict between kith, kin, and strangers: Game theory by domains. In C. Crawford & D. Krebs (Eds.), *Foundations of evolutionary psychology*. Mahwah, NJ: Erlbaum.
- Kenrick, D. T., & Sundie, J. M. (2005). How do cultural variations emerge from universal mechanisms? *Behav Brain Sci*, 28(6).
- Knobe, J. (2019). Philosophical Intuitions Are Surprisingly Robust Across Demographic Differences. *Epistemology & Philosophy of Science*, 56(2), 29-36.
- Lieder, F., & Griffiths, T. L. (forthcoming). Resource-rational analysis: understanding human cognition as the optimal use of limited computational resources. *Behav Brain Sci*.
- Macfarlan, S. J., & Quinlan, R. J. (2008). Kinship, Family, and Gender Effects in the Ultimatum Game. *Human Nature*, 19(3), 294-309.
- Machery, E. (2010). Explaining why experimental behavior varies across cultures: A missing step in “The weirdest people in the world?”. *Behavioral & Brain Sciences*, 33(2/3), 41-42.

- Machery, E., & Stich, S. (forthcoming). Demographic Differences in Philosophical Intuition: A reply to Joshua Knobe.
- Markman, A., Blok, S., Dennis, J., Goldwater, M., Kim, K., Laux, J., . . . Taylor, E. (2005). Culture and individual differences. *Behavioral and Brain Sciences*, 28, 831- 831.
- Mas-Colell, A., Whinston, M. D., & Green, J. R. (1996). *Microeconomic Theory*. Oxford: Oxford University Press.
- Massimi, M. (2018). Four kinds of perspectival truth. *Philosophy and Phenomenological Research*, 96(2), 342-359.
- Mikhail, J. (2011). *Elements of Moral Cognition: Rawls' Linguistic Analogy and the Cognitive Science of Moral and Legal Judgment*. Cambridge: Cambridge University Press.
- Mitchell, S. (2003). *Biological Complexity and Integrative Pluralism*. Cambridge: Cambridge University Press.
- Nichols, S. (2004). *Sentimental Rules: On the Natural Foundations of Moral Judgment*. Oxford: Oxford University Press.
- Okasha, S. (2006). *Evolution and the Levels of Selection*. Oxford: Oxford University Press.
- Parker, W. (2020). Model Evaluation: An Adequacy-for-Purpose View. *Philosophy of Science*, 87(3), 457-477.
- Penke, L. (2010). Bridging the Gap Between Modern Evolutionary Psychology and the Study of Individual Differences. In D. M. Buss & P. H. Hawley (Eds.), *The Evolution of Personality and Individual Differences* (pp. 243-279). Oxford: Oxford University Press.
- Piccinini, G., & Schulz, A. (2019). The Ways of Altruism. *Evolutionary Psychological Science*, 5, 58-70.

- Potochnik, A. (2010). Explanatory Independence and Epistemic Interdependence: A Case Study of the Optimality Approach. *The British Journal for the Philosophy of Science*, 61(1), 213-233.
- Rocheftort-Maranda, G. (2016). Simplicity and model selection. *European Journal for Philosophy of Science*, 6(2), 261-279.
- Royall, R. (1997). *Statistical Evidence – A Likelihood Paradigm*. Boca Raton, FL: Chapman and Hall.
- Rubin, H. (2018). The Debate over Inclusive Fitness as a Debate over Methodologies. *Philosophy of Science*, 85(1), 1-30.
- Ruiz, N., & Schulz, A. (forthcoming). Microfoundations and Methodology: A Complexity-Based Reconceptualization of the Debate. *British Journal for the Philosophy of Science*.
- Schulz, A. (2018a). By genes alone: a model selectionist argument for genetical explanations of cooperation in non-human organisms. *Biology & Philosophy*, 32, 951-967.
- Schulz, A. (2018b). *Efficient Cognition: The Evolution of Representational Decision Making*. Cambridge, MA: MIT Press.
- Schwarz, G. (1978). Estimating the Dimension of a Model. *Annals of Statistics*, 6, 461-465.
- Smith, R. J. (1992). Non-Nested Tests for Competing Models Estimated by Generalized Method of Moments. *Econometrica*, 60(4), 973-980.
- Sober, E. (1988). *Reconstructing the Past: Parsimony, Evolution, and Inference*. Cambridge, MA: MIT Press.
- Sober, E., & Wilson, D. S. (1998). *Unto Others: The Evolution and Psychology of Unselfish Behavior*. Cambridge, MA: Harvard University Press.

- Sripada, C. S. (2008). Nativism and Moral Psychology. In W. Sinnott-Armstrong (Ed.), *Moral Psychology, Volume 1: The Evolution of Morality* (pp. 319-343). Cambridge, MA: MIT Press.
- Starmans, C., & Friedman, O. (2020). Expert or Esoteric? Philosophers Attribute Knowledge Differently Than All Other Academics. *Cognitive Science*, 44(e12850).
- Stone, M. (1974). Cross-Validictory Choice and Assessment of Statistical Predictions (with Diosciussion) *Journal of the Royal Statistical Society B*, 36, 111-147.
- Stone, M. (1977). An Asumptotic Equivalence of Choice of Model by Cross-Validation and Akaike's Criterion. *Journal of the Royal Statistical Society B*, 39, 44-47.
- Thomas, M. G., Ji, T., Wu, J., He, Q., Tao, Y., & Mace, R. (2018). Kinship underlies costly cooperation in Mosuo villages. *Royal Society Open Science*, 5(2).
- Tooby, J., & Cosmides, L. (1992). The Psychological Foundations of Culture. In J. Barkow, L. Cosmides & J. Tooby (Eds.), *The adapted mind* (pp. 19-136). Oxford: Oxford University Press.
- West, S. A., El Mouden, C., & Gardner, A. (2011). Sixteen common misconceptions about the evolution of cooperation in humans. *Evolution and Human Behavior*, 32(4), 231-262.
- West, S. A., Griffin, A. S., & Gardner, A. (2007). Social semantics: altruism, cooperation, mutualism, strong reciprocity and group selection. *J Evol Biol*, 20(2), 415-432.
- West, S. A., Griffin, A. S., & Gardner, A. (2008). Social semantics: how useful has group selection been? *J Evol Biol*, 21(1), 374-385.
- White, R. (2003). The Epistemic Advantage of Prediction over Accommodation. *Mind*, 112(448), 653-683.

Zucchini, W. (2000). An Introduction to Model Selection. *Journal of Mathematical Psychology*,
44, 41-61.