Evolution and Human Behavior xxx (xxxx) xxx–xxx

Contents lists available at ScienceDirect

Evolution and Human Behavior

journal homepage: www.elsevier.com/locate/ens

Deciding what to observe: Thoughts for a post-WEIRD generation

H. Clark Barrett

UCLA Department of Anthropology, 341 Haines Hall, Los Angeles, CA, 90095-1553, United States of America

ARTICLE INFO

Keywords:
WEIRD
Evolutionary social science
Cross-cultural comparisons
Phenomics
Replication crisis
Slow science

ABSTRACT

The evolutionary social sciences (ESSs) are thriving, and seem to have entered a period of normal science. This is a good time to examine our own practices, theoretical and empirical, and to ask how we might improve. Here I review papers published in the past five years in EHB to explore major trends in the field. Theoretically, the popularity of certain topics (cooperation, mating, life history) has led to great progress, but might have narrowed our theoretical vision. Empirically, most research is still conducted in WEIRD populations, with a smaller mode of research in small-scale societies, and very little in the middle. I offer suggestions for broadening our theoretical and empirical scope, centered around the project of constructing a representative map of the human psychological and behavioral phenome.

It is the theory which decides what we can observe.
Albert Einstein, quoted in Heisenberg (1971), p. 63.

The only pressing reason for changing a theory is disagreement with facts. Discussion of incompatible facts will therefore lead to progress. Discussion of incompatible hypotheses will not. Hence, it is sound procedure to increase the number of relevant facts.
Paul Feyerabend, *Against Method* (1993), p. 26.

1. Introduction

The evolutionary social sciences (ESSs) are thriving. The controversies that plagued the field, especially sociobiology and evolutionary psychology, have largely receded. The visibility and stature of the field have increased, along with the diversity of topics studied and methods used. Evolutionary approaches in the social sciences have become more popular and even mainstream, with papers now appearing regularly in top general science journals. And ten years after Henrich, Heine, and Norenzayan's influential "WEIRD" paper (Henrich, Heine, & Norenzayan, 2010), there seems more awareness than ever of the importance of expanding our research to populations around the world.

At the same time, however, the social sciences are facing a crisis concerning the reliability, validity, and replicability of results and findings. This has surfaced most noticeably in the so-called replication crisis in psychology (Open Science Collaboration, 2015; Maxwell, Lau, & Howard, 2015; Shrout & Rodgers, 2018), arising from a proliferation of false positives, or illusory findings (Ioannidis, 2005; Simmons, Nelson, & Simonsohn, 2011; Smaldino & McElreath, 2016). This has led to much debate about the causes of the problem, including practices such as "p-hacking," "HARKing," and the "file drawer" effect (Kerr, 1998; Simonsohn, Nelson, & Simmons, 2014). Sociologically, the problem may stem from the incentive structures of science, which do not always reward truth-seeking practices (Smaldino & McElreath, 2016).

Most ESS researchers are aware of these problems. Indeed, our field might be as good or better than others at adopting emerging scientific best practices such as pre-registration and public data archiving. Despite this, however—or perhaps even partly because of it—we are still in danger of our assumptions and practices leading us astray. Because the ESSs have entered a period of what Kuhn would call "normal science" (Kuhn, 2012), now is a time to take a look at our own practices, both theoretical and empirical, and to ask how we might improve.

Unlike many critics, I do not think that there is anything especially or unusually wrong with our field. Indeed, I think we are doing as well or better than, for example, some branches of psychology. And I believe that we are now largely in a position to do evolutionary social science properly. But in order to do so, we must proceed thoughtfully. Arguably, what got psychology into its replication crisis were the rote procedures of normal science, stabilized by the comfort and group-think that come with being part of a thriving and prestigious field. Evolutionary social scientists face the same danger, if not more so, due to a widespread conviction in the field that the truth is on our side (Coyne, 2010; Segerstråle, 2000).

In this paper, I will be borrowing a page from Paul Rozin's prescient critique of the field of psychology—and in particular, social psychology (Rozin, 2001). In that paper, Rozin argues that social psychology has moved too quickly to try to become a mature or "advanced" science,

E-mail address: barrett@anthro.ucla.edu.

https://doi.org/10.1016/j.evolhumbehav.2020.05.006
Received 5 January 2020; Received in revised form 28 March 2020; Accepted 20 May 2020
1090-5138/

skipping the crucial early steps of observation—“informed curiosity” is his term for it—and moving quickly towards a hypothesis-driven, model-based experimental science. This physics-style model works well for some parts of psychology, but when applied too quickly to complex phenomena such as social cognition, it produces what some might call “the illusion of explanatory depth” (Rozenblit & Keil, 2002). Rozin argues that psychologists have largely skipped past trying to figure out what they are trying to explain—i.e., the phenomena themselves—and gone straight to theorizing. As he puts it: “Just as biologists have learned about life by studying different species and different environments, we would do well to open our eyes more widely before we dig too deep a hole at one place in the broad and varied terrain of human social life.” (Rozin, 2001, p. 13).

An important part of Rozin's argument is that the reverence for (overly) formal theory in psychology—theory that aims for a level of rigor that is unjustified given the field's understanding of the underlying phenomena, coupled with what he calls the “illusion of definitiveness in experiments”—leads to a sense of precision that conceals a wobbly underlying foundation. It can also, I suggest, lead to a proliferation of quantified bullshit: illusory claims, supported by numbers and statistics, that are difficult to tell from the truth.

Rozin's cautionary tale is one that I think the ESSs, too, would do well to heed. To what degree have we been captured by “physics envy?” How might our reverence for theory have led us astray by making us fail to see theoretical alternatives, and by herding the field into studying some questions at the expense of others? Are we looking in the right places, and in the right ways? And how might we, in Rozin's words, “open our eyes more widely”?

Before proceeding, let me reiterate that I do not believe that Rozin's critique applies to ESS to the same degree that it applies to social psychology, largely because I believe that we are on much more solid theoretical ground. Evolutionary theory provides a very strong basis for a “cumulative theoretical framework” (CTF) for human behavior (Barrett, 2015; Muthukrishna & Henrich, 2019; Tooby & Cosmides, 1992). I will argue, however, that theory-worship in our field, while largely a good thing, can have negative consequences. First, there is reverence for formalization (often without understanding it), coupled with the confidence that we and our theoretical ancestors are doing it right. For consumers of theory, as most ESS researchers are, this often occurs without enough careful scrutiny of the many degrees of freedom between evolutionary models and actual behavior. Second, our love of theory leads to a rush to test it, often by turning theory into methods in a hasty way: for example, by turning a model parameter into a laboratory prime, even if theory says nothing about whether it should work that way. Finally, we have begun to settle too much into what I'm calling “rote procedures of normal science,” including how we select questions, methods, and study populations, all of which are still based more on convenience and less on principle than most of us would like to believe. I will provide examples and evidence for these claims below.

This kind of normalcy is a recipe for, at best, disciplinary stasis, and at worst, a skewed view of human nature—just as Rozin argued for psychology. Coupled with the self-confidence that ESS practitioners have about the truthiness of evolutionary theory are the incentive structures of “big science” that push us towards research that is increasingly bigger, faster, and shinier (Smaldino & McElreath, 2016). Against this backdrop, I will argue that at least some of the field would benefit from a strategy of slowing down, loosening up, and looking more carefully at the scope of human phenomena we seek to understand. Beginning with a brief survey of the current state of the field, I will ask how our assumptions and practices may be narrowing our view of human nature, and how we might open our eyes more widely.

2. A brief sketch of the field

Before you read on, ask yourself how you think we are doing. Is our view of humans and human nature as broad and comprehensive as it could be? If not, why not? Are we focusing too obsessively on some questions or phenomena at the expense of others? And, ten years after the WEIRD paper, how representative is our research of our species as a whole?

To get some empirical traction on these questions, I conducted a bibliometric review of all papers published in this journal in the last five years (2015 to 2020). I downloaded data on all 300 articles published in EHB during this period from the Web of Science platform, and analyzed them in R (see Supplementary Material for details). Obviously, this is not the same as a comprehensive analysis of all of evolutionary social science, because much work, including some of the highest-profile work in the field, is published outside of EHB. Still, as the flagship journal of the field, one would hope that it provides a reasonably representative snapshot (for a related, in-depth analysis with similar conclusions see (Pollet & Saxton, 2019)).

My analysis focused mainly on two questions, theoretical and empirical. The first question is about *what* we study, and how it is informed by theory. Again, my question is not whether the ESSs are based on a solid theoretical foundation; I think they are. Instead, I want to ask how our theories might be biasing us to look at only some aspects of human nature, and perhaps to look at them in ways that lead to a skewed or incomplete understanding. The second question is about *who* we study, and how this might influence our conclusions about human nature as well. This, of course, was the primary concern raised in Henrich et al.'s (2010) WEIRD paper, and it is worth asking, ten years on, how we are doing.

2.1. Theoretical trends

A sense of the theoretical proclivities of our field can be gleaned from the keywords that authors provide for their papers. A list of the top 20 keywords and their frequencies are shown in Table 1 (I have removed the keywords “evolution” and “evolutionary psychology,” which were near the top of the list).

At the top of the list are, arguably, the four currently most popular theoretical domains of research in the ESSs. These are (1) cooperation, (2) sexual selection, (3) life history theory, and (4) cultural evolution. Interestingly, each of these is accompanied by a body of formal theory, including mathematical theory, that can be used to derive and test predictions. As I'll argue, this is both good and bad. It's mostly good, because it means that work in ESS is grounded in formal theory to a greater degree than in, for example, psychology. However, it also has downsides. It means that work in the field is heavily focused on these

Table 1
Most frequent keywords in EHB, 2015-2020.

Keyword	Frequency
Cooperation	26
Sexual selection	16
Life history theory	14
Cultural evolution	14
Mate choice	12
Attractiveness	11
Dominance	9
Mate preferences	9
Sex differences	9
Social learning	8
Competition	7
Hunter gatherers	7
Masculinity	7
Attraction	6
Emotion	6
Kinship	6
Life history strategy	6
Marriage	6
Partner choice	6
Disgust	5

areas, at the possible expense of other aspects of human cognition and behavior. It also means that our reverence for theoretical work in these areas might lead us to assume that the theoretical framings we've chosen are the *right* ones. I'll give some examples below of how this might have blinded us to other theoretical possibilities which, in retrospect, seem like somebody should have noticed.

2.2. Empirical representativeness

The second goal of my literature review was to ask how representative our research is of the human species, in terms of the people we study. To assess this, I focused only on papers that gathered original human subjects data (excluding, for example, meta-analyses, archival, and non-human primate studies). I categorized study populations to nation level, and whether or not they included "small-scale" societies (see Supplementary Materials for details).

The large majority of human subjects research papers published in EHB do not, even now, report the country of origin of their study participants in the abstract. When not mentioned, these samples are typically student samples (largely in the US but also frequently in Europe or east Asia), or mTurk. For this analysis, I coded such samples as "generic samples" (GS). These are "generic" in the sense that data are presented as if participants' cultural identities do not matter for the conclusions being drawn.

Table 2 and Fig. 1 summarize the results of this analysis. Of 228 papers that gathered original human subjects data published in EHB between 2015 and 2020, 72% used generic samples, 12% involved cross-cultural comparisons, and 14% included research in small-scale societies (see Supplementary Material for details).

Of the 228 papers with original human subjects data, 86 used non-generic samples. Of these 86 studies, 59 used samples from a single country, and the remaining 27 involved cross-cultural comparisons (two studies sampling a very large number of countries were excluded from the map; see Supplementary Materials for details). Fig. 1 shows a map of the countries from which the 86 non-generic samples were drawn, coded by frequency of sampling from a given country (note that participants in many of these cases were college students, but their nationalities were identified in the abstract).

There is much to be admired in this snapshot of ESS research, including our efforts to study non-WEIRD people. However, it also suggests that the majority of our studies continue to be based on "convenience," in several senses. First, by far the most common participant in ESS research is a college student, or an mTurker. Second, even cross-cultural work often involves "convenience sampling," as opposed to theoretically motivated sampling: i.e., selecting sites based on one's academic network (something I do myself, and better than just using college students, but still a form of convenience rather than theoretically principled sampling). For many studies run in small-scale societies, there is often no more justification than "this has never been studied in non-WEIRD people." It could be argued that just as college students are treated as "generic" for certain kinds of questions, so are people in small-scale societies. And there is a large, unsampled middle of people on earth who are neither (1) college students nor (2) subsistence-level villagers. When it comes to coverage of the human species, we can do better.

I will now turn to my central arguments. Empirically, the snapshot just presented suggests that the "normal science" of our field, including our particular theoretical preferences and methodological habits, may have narrowed our view of human nature in ways that we don't yet know. I will argue that our field can benefit from loosening up, looking at a greater range of phenomena through wider lenses, without losing the sense of rigor and precision that our field aspires to. By doing so, I suggest, we can produce a more comprehensive and accurate picture of the phenomenon we are trying to describe: human nature.

3. Loosening up

The two quotes at the beginning of the paper, by Einstein and Feyerabend, illustrate a tension between two approaches to research alluded to by Rozin. At Einstein's end is theory-driven observation, which I'll call the classical or "physics" model of research. At Feyerabend's end is the method often advocated by ethnographers: informed curiosity, or "just looking." Feyerabend, a scientific anarchist, at times seemed to be making a quasi-Darwinian argument for the importance of variation-generating processes in driving knowledge production and dislodging us from sub-optimal peaks in knowledge space.

Few philosophers of science would draw a hard line across the continuum between theory-driven and exploratory research, and there is a gamut of views regarding how theory does, and/or should, inform observation (Godfrey Smith, 2003). That said, my impression is that most ESS researchers orient more towards Einstein than Feyerabend. As evidenced by the reactions of several reviewers to an earlier draft of this paper, it's clear that ESS researchers view the theoretical grounding of our field as perhaps its strongest point. On this I agree: theory is good. Without it, it's hard to know what science would be.

On the other hand, there are pejorative terms for exactly the same orientation. "Physics envy" is sometimes used to describe a desire for theoretical and methodological rigor beyond what a field (currently) merits. Rozin persuasively argued that such an orientation has caused harm in psychology by lending an air of rigor and definitiveness to phenomena that are, in fact, far less well understood than theory would suggest. I want to tread carefully here: in my view, evolutionary theory is quite rigorous, and "cognitive" theory (e.g., computationalism) also, to a degree; "social" theory, perhaps, a bit less so. There is no question, in my mind, that the ESSs should be grounded in these bodies of theory, and there are many good examples of work unifying evolutionary, developmental, cognitive, and behavioral levels of causation and explanation.

That said, however, I would argue that theory worship—the uncritical adoption of theory, or tendency to take it on faith—can sometimes constrain thinking in ESS, making us more theoretically conservative than, for example, biologists can be (as in the slow adoption of multi-level selection theory; see below). Theory worship can fuel "theory tenacity," the ability of theories to hold on despite empirical or theoretical reasons to discard them ((Gowaty, 2018); see Loehle (1987) for an argument similar to Rozin's in the field of ecology). Our admiration for the beauty of evolutionary theories might be normal—especially in Kuhn's sense—but it merits reflection. At its best, theory

Table 2
Study types and sample populations in EHB, 2015-2020.

Study type	N papers	No small-scale sample	Includes small-scale sample	Generic sample only	One country only
Original human subjects data	228	196 (86%)	32 (14%)	142 (72%)	200 (88%)
Database	18	7 (39%)	11 (69%)		
Large dataset (N = 45,87)	2	2	0		
Metaanalysis	6				
Non-human subjects	7				
Non-empirical	39				
Total	**300**				

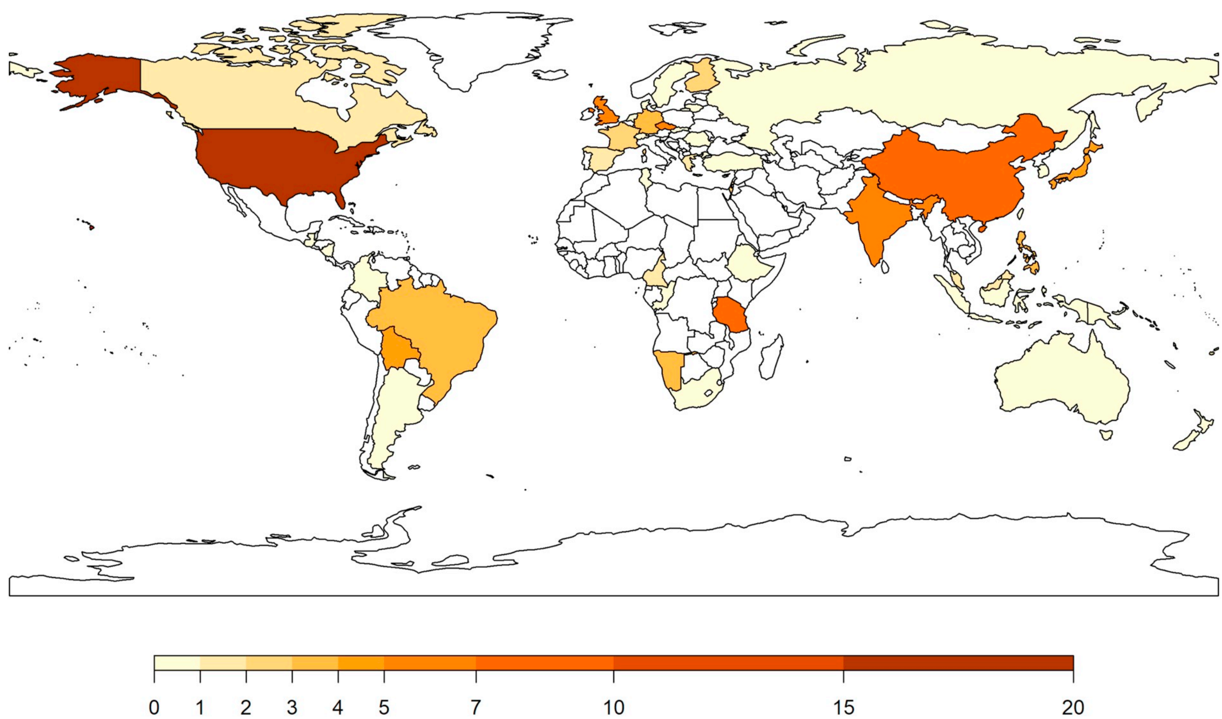

Fig. 1. Country representation in non-generic samples (N = 86 studies).

leads to progress in our understanding of the evolution of human nature by telling us where to look. At its worst, however, theory may blinker us, for exactly the same reason.

Let me briefly give examples from three top theoretical domains of research in our field (cooperation, sexual selection, and life history theory) in which the trajectory of research has been shaped by theoretical convictions that turned out to be narrower than they could have been, creating historical path dependencies that continue to be felt today.

3.1. Multi-level selection theory

In the early days of the ESSs, it seems safe to say that the idea of group selection was regarded as a terrible mistake. I, for one, was led to believe that it simply could not be true on a priori grounds (similarly, I suppose, to what we were taught about Lamarckism, though that was held to be just empirically false and never to be entertained). This message was hammered home by the persuasive verbal arguments of Williams (2018) and Dawkins (2016). Even at the time, more mathematically sophisticated practitioners such as (Price, 1970, 1972) contradicted this message (emphasizing the importance of proper training in formal theory). But the point is that for many ESS researchers then and today, group selection, and multi-level selection more generally, are regarded as either theoretically impossible or at best a form of semantic gamesmanship (Pinker, 2015; West, Griffin, & Gardner, 2008). Many theoretical biologists would disagree (Okasha, 2006; Traulsen & Nowak, 2006).

Anti-group-selectionism has had a noticeable influence on the design of research and the interpretation of findings in ESS. For example, reciprocal altruism, which can be seen as an individual-level mechanism of cooperation, was and is regarded by many as the best, or perhaps only, possible explanation for human cooperation, even in one-shot encounters (Delton, Krasnow, Cosmides, & Tooby, 2011). However, other mechanisms, such as indirect reciprocity and cultural group selection, are theoretically capable of explaining this kind of behavior (Boyd & Richerson, 1990, 2010; Leimar & Hammerstein, 2001; Panchanathan & Boyd, 2004). A large number of empirical studies, including studies with economic games, have been performed in an effort to adjudicate between these models (Burton-Chellew & West, 2013; Debove, Baumard, & Andre, 2016; Henrich et al., 2001; Rand, Tarnita, Ohtsuki, & Nowak, 2013). Surprisingly, however, despite the remarkable precision of formal theory in this domain, there remains a stalemate between individual and multilevel selectionist camps in how to explain findings (Burton-Chellew & West, 2013; Delton et al., 2011). The process does not look terribly Popperian. Theory-worship and theory tenacity seem at least as explanatory of the inferences that many ESS researchers draw about human nature in this domain.

The solution here is not "less theory"—that is not what I am proposing. It is more theoretical *openness*, including to the possibility of theories not yet on the table. After all, what is "impossible" in a model might occur right before your eyes in real life. Physics provides some of the most interesting examples.

3.2. Parental investment theory

A second example concerns the use of sexual selection theory in the ESSs, particularly in evolutionary psychology. The theory of sexual selection was of course pioneered by Darwin, but the incarnation that has come to dominate ESS research is Trivers' parental investment theory, influenced by the work of Bateman (Bateman, 1948; Trivers, 1972). Alongside the primacy of individual selection, many of us were taught in graduate school that sex differences flow directly and mechanistically from differences in parental investment—the differential cost of reproduction for males and females—which in turn follow from anisogamy (Parker, Smith, & Baker, 1972). It is difficult to overstate the degree to which research on mating and sex differences in the ESSs have been shaped by this theoretical perspective, which seems to show on nearly a priori grounds, barring special conditions of sex role reversal, that females must be choosier than males.

Surprisingly for some, theorists have recently begun to call into question the theoretical assumptions of parental investment theory, as well as its empirical basis. For example, (Kokko & Jennions, 2008) noted that parental investment theory, as originally formulated, overlooked frequency-dependent selection on investment due to the fact that each offspring has exactly one male and one female parent (increasing, not decreasing, selection for investment by the more competitive sex). Mathematically, they show that the Trivers / Bateman scenario involves special assumptions that were taken as given in the original theory, and discuss how these assumptions influenced future research by becoming baked into models (see also (Kokko, Jennions, & Brooks, 2006; Queller, 1997).

Parallel to this, at least some of the empirical foundations of the Bateman / Trivers model have begun to erode. In perhaps the most startling example, Patty Gowaty and colleagues have shown, both through a re-analysis of Bateman's data and a replication of his experiments, that Bateman's classic findings about reproductive skew in fruit flies—*the* grounding example of parental investment theory in textbooks—were false (Gowaty, Kim, & Anderson, 2012). To add insult to injury, Gowaty points out that this could have and should have been noticed long before, because it involved a simple mathematical mistake. She invokes the idea of theory tenacity to explain why it took so long for anyone to notice (Gowaty, 2018). In the human case, as well, cross-cultural work is beginning to call into question the idea that sex differences in choosiness are a universal feature of human nature (Scelza et al., 2020).

3.3. Life history theory

Life history theory (LHT) is a fast-growing area of research in the ESSs, with studies examining the “fast / slow” continuum in humans becoming particularly popular (Figueredo, Vasquez, Brumbach, & Schneider, 2007; Griskevicius, Tybur, Delton, & Robertson, 2011). These are early days, but it seems possible that there is a rush to design studies and explain findings before the theory is fully baked. Recently several theorists have raised concerns about how LHT is being used, particularly in evolutionary psychology (Stearns & Rodrigues, 2020; Zietsch & Sidari, 2019). Stearns & Rodrigues (2020) point out that the fast-slow continuum was originally proposed to explain phylogenetic patterns between taxa and not within species, where the evidence is scarcer. They also argue that assumptions about fast-slow plasticity in humans are poorly theorized, both in terms of conditions (e.g., environmental harshness, psychosocial stress) and outcomes (e.g., psychological traits). In contrast to this, many ESS researchers seem to think this is a well-theorized area with clear predictions, and the theory is being used to explain a wide variety of phenomena, from psychological responses to poverty to shifting behavioral decisions in response to laboratory primes (Sng, Neuberg, Varnum, & Kenrick, 2017).

My goal is not to chastise researchers in these areas for relying on existing theory, or working with it. Instead, I am using these examples to show how we can, sometimes, be blinkered by theory, and by assuming that smart people before us have set up models correctly. Against this, loosening up, in the sense of considering alternatives and insisting on interrogating the logic of our theories, is at least one counter-strategy. My experience is that ESS researchers are frequently conversant with formal mathematical models, but often reduce these quickly to verbal arguments that have many hidden assumptions—and these become baked into our thinking (Kokko & Jennions, 2008; Kokko et al., 2006). If Einstein is right that it's the theory that decides what we observe, this can be problematic. Theory tenacity can cause us to take data as confirming of a theory even when there are other options. For example, LHT is frequently used to explain behavioral responses to poverty, such as risky behavior. But is behavior such as drug use and crime really best explained by an evolved reaction norm, or are there other explanations we should consider?

4. Looking carefully

While theories influence what we observe, it is ultimately we who decide. So what is it that we *want* to observe? There is no single goal or set of goals of ESS research, but I would argue that two broad goals of the field are to understand human universals and human variation, both of which I place under the rubric of human nature. A proper approach to human nature must consider (1) the features that we share with other species due to our deep common ancestry, (2) derived traits in our lineage that explain uniquely human aspects of cognition and behavior, and (3) the ways in which humans vary as a function of history, culture, environment, and individual experience—all of which are flexibility enabled by our biology (Barrett, 2015).

4.1. Mapping the elephant

Using the popular metaphor of the elephant (the blind men and the elephant, the elephant in the room), let's call this sum total of what we want to know “the elephant.” We are trying, metaphorically, to assemble a picture, portrait, or map of it. The map must, of course, be smaller and less detailed than the territory itself, but a good map is both *accurate* and *complete*, neither falsely representing nor leaving out parts of the territory. Of course some maps, such as highway and subway maps, privilege certain useful features at the expense of others, but can still be accurate representations of portions of the world, given their representational conventions. Ultimately it could turn out that for understanding human evolution some features of the elephant are more important than others, and should be examined and delineated in greater detail. Until we know what those are, however, it seems that an accurate, complete, and representative portrait of human nature should be our goal. We can call this goal a map of the human psychological *phenome*—the sum total of our psychological, behavioral and cultural traits, and how they vary, or do not, across people (Houle, Govindaraju, & Omholt, 2010).

This is, of course, a tall order: a huge elephant, requiring a giant canvas, and many painters. We don't even know, at this point, entirely what such a map would look like. But we can decide if such a map is one of our goals, and if it is, how suitable our current theory and methods are for helping us construct it.

The sketch of the field I presented above gives some sense of how our efforts are allocated, theoretically and empirically, but does not tell us how representative our work is of the elephant itself. Cooperation is clearly important to the evolution of our lineage and our psychology—but how important? Similarly, sexual selection is undoubtedly important for shaping some of our phenome—but how much? Are the top four topics in Table 1 (or the top 10, 20, or N) likely to be adequate for constructing our map?

It might be argued that some of the theoretical and empirical skew of our field, the “light is better here” feel of our preference for certain phenomena, is justified. For example, one reason we might want to systematically map some human traits in more detail than others might be in the search for *uniquely human* traits—autapomorphies, in the jargon of phylogenetics. The vast majority of human traits are not uniquely derived in our lineage, and are shared homologously with other primates, mammals, vertebrates, etc. For good reason, much effort has been devoted to trying to understand the unique features of our lineage, and the evolutionary processes that shaped them.

4.2. The search for uniquely human traits

In the domain of cognitive traits, I recently surveyed the literature on human cognitive specializations, looking for domains in which there is evidence for uniquely derived mechanisms or abilities in humans (Barrett, 2017). There is reasonable evidence for human-specific cognitive specializations in (at least) the following domains: (1) language, (2) mindreading, (3) cultural transmission, (4) cooperation, (5) tool use, (6) executive functions. In each of these domains, in turn, there is evidence for sub-mechanisms (for example, human-specific language abilities utilize derived mechanisms of both perception and production). This is not an exhaustive list, and our understanding of human autapomorphies is in its early days. But it gives a snapshot of areas in which humans may be significantly different, psychologically and behaviorally, than our closest relatives. And it is worth noting that most current "big picture" attempts to account for the evolution of uniquely human cognition and behavior—what makes us special—draw from somewhere on this list: e.g., Deacon's *The symbolic species,* Richerson and Boyd's *Not by genes alone: How culture transformed human evolution,* and Tomasello's *The cultural origins of human cognition* (Deacon, 1997; Richerson & Boyd, 2008; Tomasello, 2009). Accounts such as these, while not purporting to explain every last detail of human nature, are certainly intended as theories of the elephant.

That said, it is worth asking just how good our current picture of the elephant is, both theoretically and empirically. Empirically, have we adequately captured and described human nature, enough to properly understand what it is that we're trying to explain? And theoretically, do we have the right account of how human nature evolved—or even a set of accounts that is complete enough to include what may one day prove to be the right one?

I doubt that anyone would be bold enough to suggest either. And because our field is still young, we obviously have a long way to go. But here we might well consider Rozin's advice, which is not to "dig in" at the spot where the field happens to find itself, either theoretically or empirically. There is still a wide gap between our understanding of uniquely human traits—which is itself patchy—and our ability to satisfyingly explain them using our current evolutionary toolkit. As an example, let me briefly consider our ability to explain the evolution of a very well-studied, but still poorly understood, human trait: theory of mind.

4.3. The theory / phenomenon gap: theory of mind as an example

The ability known as "theory of mind" (ToM), or mindreading, is a good example of a psychological ability that is uniquely elaborated in humans. This complex ability is difficult to describe succinctly, but it can be summarized as our sensitivity to others' minds: our abilities to perceive, infer, and respond to others' mental states (desires, emotions, moods, beliefs, knowledge, intentions, etc.) (Apperly, 2010; Nichols & Stich, 2003; Wellman, 1992). Empirically, it seems well established that humans have uniquely derived abilities in this domain (Adolphs, 2009; Saxe, 2006). We're not the only animals with theory of mind, but, consider this analogy. Hummingbirds aren't the only animals that can fly, yet their skills in this domain are unique. As aerial agility is to hummingbirds, theory of mind is to humans.

But why? Why are we so good at this? What, exactly, does this ability consist of, how is it used in real life, and what are the fitness benefits that drove its evolution?

I'd argue that despite thousands of empirical papers on theory of mind, including many developmental studies on tasks such as the false belief task, we're still not in a position to give a satisfying answer to these questions. Take the question of function. Proximately, the function of mindreading seems to follow from its description: it allows us to make use of the mental states of others to predict their behavior and understand their actions (e.g., what that person gesturing at you is trying to communicate, what the person struggling with shopping bags by their front door is trying to do; for more on the richness of the ability see Barrett, 2015). Ultimately, this likely provides fitness benefits of various kinds: some related to Machiavellian intelligence (cooperation, deception, manipulation, social strategy), some related to improved communication and information transfer, e.g., via language, some related to social learning, e.g., through teaching, and probably more besides (Csibra & Gergely, 2011; Humphrey, 1976; Sperber & Wilson, 1995; Whiten & Byrne, 1997). But just because we can see these possible connections does not mean that we are in a position to build or test a rigorous theory of how and why ToM became elaborated as it did in humans. Nor, in fact, do we have a full empirical grasp of what ToM *is,* stemming largely from the fact that we study it using experimental tasks and not as it is deployed in the real life (Bloom & German, 2000). There are some naturalistic studies, but our phenomic understanding of ToM is largely limited to highly detailed lab tests of tiny slices of the ability (false belief tracking, emotion reading, and some others).

For explaining phenomic traits like ToM, it's not that we don't *necessarily* have good theory; but there is still a huge gap between our theoretical tools, such as they are, and what we are trying to explain. Following Feyerabend, we can theorize all we want, but until we have sufficient data, it won't make much difference in the hill-climb towards genuine understanding. At best, current pronouncements that we understand the evolutionary story of ToM—such as Heyes' (2018) claim that it is a culturally installed "cognitive gadget"—seem highly premature, and aspire to a rigor and definitiveness beyond what current theory and data can allow. And this is the case for one of the most highly studied, flagship examples of human nature. How much worse must it be for phenomena that have received a smaller share of our research effort?

4.4. How we look

Another way in which our view of the elephant may be too narrow is empirical, stemming from what we choose to observe, and who we choose to observe it in.

The modal research project in our field is lab-oriented, and experiment based. Again, this is not meant as a totalizing claim: there is plenty of methodological diversity in our field, including a lot of field work involving behavioral observation and other kinds of data besides. But the center of research gravity in the field certainly remains psychology-style studies, with most prestige attached to experiments and other kinds of quantifiable measures such as scales and biometrics.

What's the matter with this? In some ways, nothing, if these methods are used properly. But if we are asking how our methodological orientations might be narrowing our view, the focus on experiments and quantification might have downsides. For one, there is what Rozin called "the illusion of definitiveness in experiments." Experiments do allow for control, and therefore a degree of rigor, but as the replication crisis has revealed, many seemingly precise findings may in fact be quantified bullshit. The fallout from the crisis is giving us some valuable lessons about when experiments work well in producing knowledge and when they don't. By some measures, half or more of psychology studies seem to replicate, but this seems to depend mostly on the phenomenon under study (Klein et al., 2018). An important question is to what degree we are focused on the phenomena we seek to observe, as opposed to fetishizing particular methods for measuring them (such as, e.g., scales as a measure of psychological differences). As psychologist Lee Cronbach observed, "An investigation carried out under reproducible but highly specific conditions is not usually of great value in itself" (Cronbach, 1986, p. 95). We do not yet know how much of our view of human nature suffers from Cronbach's problem, the precise measurement of tiny regions of the elephant.

Experiments are, typically, pre-designed to test a particular hypothesis or hypothesis. Here Einstein's dictum applies in its strongest form, because experiments usually can't tell you anything but what you want to know (and sometimes, not even that, if theories about the underlying constructs and mechanisms are not well-specified). Again, this implies a certain circularity. If, in Rozin's terms, we've decided where to dig a hole, then experiments may be a good way to go, but they are not necessarily ideal for surveying the landscape.

Our view of the elephant is also limited by our choice of research populations, and how we pair research questions with the people and places we study them. As the analysis above reveals, our research is highly bimodal. Most ESS research is done in generic college student samples in the U.S., Europe, China and Japan. The second mode is in small-scale societies such as Hadza and Tsimane. Again, there is nothing in principle wrong with this, and work in such societies is an important antidote to the WEIRD problem. We mark small-scale societies as special for evolutionary work, perhaps because of (sometimes questionable) assumptions about the relevance of their lifeways for testing evolutionary hypotheses (Marlowe, 2005). But should our portrait of human nature be mostly based on this union of disjoint sets: college students and subsistence-level villagers?

Added to this, our pairing of research questions with populations is much more unsystematic and patchy than we might imagine. For a long time, certain questions were mostly studied in small-scale societies and not elsewhere (e.g., foraging, food sharing). Other topics, such as mate choice, personality, and politics, tended to be studied mostly in college students. This has begun to change, especially in the direction of bringing previously WEIRD studies outside the lab. However, it's still the case that many researchers, when seeking a population for a non-WEIRD replication of a phenomenon, look to small-scale societies, which may be unrepresentative in their own ways of humans more generally. Many papers in the literature review above offered no specific justification for why a study was run in a particular small-scale society.

An alternative, of course, is systematic pairing of research questions with research populations. Here it is useful to widen our lenses and realize that all human populations have characteristics, beyond just being human, that can be useful for testing evolutionary hypotheses (Apicella & Barrett, 2016). Relatively few studies in the literature review seemed to sample populations systematically based on features specifically relevant for the phenomenon under investigation, though there were some noteworthy examples. For example, Thornborrow, Jucker, Boothroyd, & Tovee (2018) used existing differences in television exposure across different regions of Nicaragua to examine how attractiveness preferences are shaped by media exposure, important for understanding the evolutionary basis of attractiveness judgments. There is much room for such inventive use of human variation that ESS researchers have yet to explore.

4.5. What we look at

A final way that we may be blinkering ourselves, in the manner Rozin suggests, is in the scope of phenomena we examine. Part of this comes from trends and fashions in the field, such as the huge outpouring of research on mate choice, or the current wave of work on LHT. Some of this is theoretically motivated, which can be good. But following trends might not, for reasons explored above, be the best strategy for a thorough map of the elephant. And some of our theoretical proclivities, I think, may stem from our intuitions about what phenomena count as “evolutionary”. To some degree, we may essentialize what counts as evolutionary by imagining whether or not it was a domain or phenomenon that existed in the evolutionary past.

Foraging and mate choice seem inherently evolutionary; cell phone use seems less so. Ditto for lots of contemporary but important phenomena in politics, religion, and social life. But there are potentially two mistakes here: (1) thinking that some supposedly “evolutionary” phenomena aren't influenced by history, technology, and modern institutions, and (2) thinking that there are some things that humans do that aren't evolutionary phenomena. Of course, much ESS research does look at historically recent phenomena, such as online social networks and internet dating, but this is often because these are seen as convenient ways of accessing evolved psychology (e.g., mate choice, status-seeking), rather than as interesting evolutionary phenomena themselves. The phenomena of contemporary life are more than just technological lenses through which to view innate preferences.

Because everything about humans is the product of evolution, all human activities are “evolutionary” and worthy of study in the ESSs—especially if they occupy major portions of our psychological and behavioral space (nationalism, globalization, environmentalism, migration, racism, etc.). There is a disjuncture, topic-wise, between what is studied in other social sciences, such as sociology and sociocultural anthropology, and what is studied in the ESSs. In the interest of genuinely understanding the human phenome, we should ask why there are major phenomena that others find important and that we, as a field, largely ignore.

5. Thinking phenomically

The same year that Henrich et al. published their WEIRD paper, biologist David Houle and colleagues published a paper called “Phenomics: The next challenge”, in which they argued:

> Phenotypic variation is produced through a complex web of interactions between genotype and environment, and such a ‘genotype–phenotype’ map is inaccessible without the detailed phenotypic data that allow these interactions to be studied. Despite this need, our ability to characterize phenomes — the full set of phenotypes of an individual — lags behind our ability to characterize genomes.
>
> Houle et al., 2010

Currently, the phenomics of plants is a thriving field (Tardieu, Cabrera-Bosquet, Pridmore, & Bennett, 2017), but there is as of yet no phenomics of humans—with the exception of work in the neurosciences (Bilder et al., 2009). There is no phenomics of most everyday human cognition and behavior. We, the ESSs, are in perhaps the best position to create one.

Unfortunately there is no algorithmic recipe for doing so, but I suggest that many of the problems alluded to in this paper could be alleviated, at least in part, if more of our field explicitly viewed itself as collaborating to produce, as a product or byproduct of our work, a map of the human phenome. This map could be used to test limitless hypotheses, but need not be collected, as most of our data currently are, only in the service of investigating a particular theory.

As a point of comparison, consider HRAF. Researchers use this quite explicitly as a kind of phenome of traditional societies, and the fact that it is (allegedly) “theory-neutral,” comprising ethnographic data not collected in the service of the uses to which it is now put, is regarded as a virtue. Whether HRAF is, in fact, a representative map of the elephant is another question; but ESS researchers use it as such.

Now imagine how much better it would be if we had a *proper* phenomic map of the human species, a dynamic and ever-growing one, one not as dependent on the specific proclivities of past ethnographers regarding what to observe (e.g., the “traditional” topics of ethnography such as kinship and marriage). I suggest that we have the tools right now to work towards this goal by altering our practices in certain ways, some of which are consonant with more general trends in best practices of science.

Practices such as pre-registration, public data archiving, and publication of null results are all aimed at reducing scientific bias or skew due to our theoretical predilections and pet hypotheses. They allow later researchers to use our data for different purposes, to remove (to some degree) the theoretical spin we've placed on our findings.

In the same vein, I suggest the following practices that can allow us to begin constructing a proper map of the human psychological and behavioral phenome, while at the same time continuing to test hypotheses:

1. When a study is designed to test a specific hypothesis, specify it (already required in pre-registration). When possible, in pre-registration materials or the paper itself, discuss ways in which the theoretical lens might have influenced observation; for example, some scales or instruments may be designed explicitly to maximize variation in a particular dimension, rather than looking for similarities. In ToM research, for example, tests such as the “mind in the eyes” task were explicitly designed to look for individual differences in emotion-reading, and differ from other tasks such as the false belief task, which is a blunt measure of individual differences in ToM abilities (Baron-Cohen, Wheelwright, Hill, Raste, & Plumb, 2001).
2. For every hypothesis-oriented study, attempt to design the data collection protocol so that the data could, in principle, be used by

other researchers for purposes other than just testing or re-evaluating your hypothesis. For example, consider including general-purpose measures in addition to more boutique, hypothesis-specific measures. As a field, we need to work towards better ways to standardize data reporting, e.g., via metadata tags, to allow construction of phenomic databases. Other fields, such as neuroscience, are working towards such standardized data formats specifically to allow for post-hoc testing of hypotheses using previously collected data (Parker et al., 2008). This is an area where the ESSs could innovate.

3. Make clear the rationale for population selection. In the case of cross-cultural comparisons, if the rationale is simply a WEIRD vs non-WEIRD comparison, state whether the selection of populations was principled, or based on convenience (e.g., snowball sampling of sites via your personal academic network). Consider representativeness, both statistical and theoretical: Make explicit how your research participants were, or were not, selected representatively for your phenomena of study. When possible, study a phenomenon in a particular population for a reason; if not, say so.
4. In the abstract (and if possible, title) of your paper, state clearly who the participants in your study were (e.g., "in a sample of US college students, we showed…"). Journals, including EHB, should consider requiring metadata tags with participant demographic information, to facilitate the construction of searchable phenomic databases.
5. When possible, collect demographic data from participants, even "generic" participants, that could allow them to be compared to other populations; e.g., SES, language proficiency, religiosity, political orientation, cultural heritage, etc.
6. To reduce the "bimodality" problem, consider studying participants in the large middle ground between WEIRD college students and subsistence-level societies, when they are just as appropriate or more so for studying your target phenomenon.
7. To reduce the essentializing of evolutionary topics, study, and allow students to study, topics that have not typically been examined through an evolutionary lens. Do not require evolutionary hypotheses going in; allow exploratory research to capture dimensions of a phenomenon that could prove to have evolutionary relevance. Many of our hypotheses are no different than hunches anyway, so hunches should count in deciding where to look.

6. Conclusion

My goals in this paper have been to offer a critical examination of the current state of our field, and positive recommendations for how we might improve in our shared goal of understanding the evolution of human behavior. The ESSs are thriving, but settling too much into a normal science mode risks stasis. A careful look at where our field stands can help us to see where there is room for improvement.

Theoretically, while our field has impressive theoretical grounding and breadth, there are theoretical trends that lead us to study some topic over others. It is not clear whether our current set of theories will allow us to fully understand and describe the evolution of human cognition as a complex whole, as opposed to a disjoint set of phenomena (e.g. mating, cooperation, language). Empirically, our research is still much less culturally and thematically diverse than it could be, with a bimodal distribution of research participants. We should expand our work to cover the rest of humanity, and attempt to describe human diversity and similarity through a phenomic lens.

Ultimately, it is we who decide what to observe. By slowing down and looking in new ways, we can develop a better understanding of us as a species, and how and why we are the way we are today.

Funding

This research was supported by the Geography of Philosophy Project, funded by the John Templeton Foundation (60813).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.evolhumbehav.2020.05.006.

References

Adolphs, R. (2009). The social brain: Neural basis of social knowledge. *Annual Review of Psychology, 60*, 693–716.
Apicella, C. L., & Barrett, H. C. (2016). Cross-cultural evolutionary psychology. *Current Opinion in Psychology, 7*, 92–97.
Apperly, I. (2010). *Mindreaders: The cognitive basis of "theory of mind"* Psychology Press.
Baron-Cohen, S., Wheelwright, S., Hill, J., Raste, Y., & Plumb, I. (2001). The "Reading the Mind in the Eyes" Test revised version: A study with normal adults, and adults with Asperger syndrome or high-functioning autism. *The Journal of Child Psychology and Psychiatry and Allied Disciplines, 42*(2), 241–251.
Barrett, H. C. (2015). *The shape of thought: How mental adaptations evolve.* Oxford University Press.
Barrett, H. C. (2017). The search for human cognitive specializations. *Evolution of nervous systems*(2nd ed.). *Vol. 4. Evolution of nervous systems* (pp. 355–366). Elsevier.
Bateman, A. J. (1948). Intra-sexual selection in Drosophila. *Heredity, 2*(3), 349–368.
Bilder, R. M., Sabb, F. W., Cannon, T. D., London, E. D., Jentsch, J. D., Parker, D. S., … Freimer, N. B. (2009). Phenomics: The systematic study of phenotypes on a genome-wide scale. *Neuroscience, 164*, 30–42.
Bloom, P., & German, T. P. (2000). Two reasons to abandon the false belief task as a test of theory of mind. *Cognition, 77*(1), B25–B31.
Boyd, R., & Richerson, P. J. (1990). Group selection among alternative evolutionarily stable strategies. *Journal of Theoretical Biology, 145*(3), 331–342.
Boyd, R., & Richerson, P. J. (2010). Transmission coupling mechanisms: Cultural group selection. *Philosophical Transactions of the Royal Society of London B: Biological Sciences, 365*, 3787–3795.
Burton-Chellew, M. N., & West, S. A. (2013). Prosocial preferences do not explain human cooperation in public-goods games. *Proceedings of the National Academy of Sciences, 110*(1), 216–221.
Open Science Collaboration. (2015). Estimating the reproducibility of psychological science. *Science, 349*(6251), aac4716.
Coyne, J. A. (2010). *Why evolution is true.* Oxford University Press.
Cronbach, L. J. (1986). Social inquiry by and for earthlings. *Metatheory in social science: Pluralisms and subjectivities* (pp. 83–107). .
Csibra, G., & Gergely, G. (2011). Natural pedagogy as evolutionary adaptation. *Philosophical Transactions of the Royal Society B: Biological Sciences, 366*, 1149–1157.
Dawkins, R. (2016). *The selfish gene.* Oxford University Press.
Deacon, T. W. (1997). *The symbolic species: The co-evolution of language and the brain.* Norton.
Debove, S., Baumard, N., & Andre, J.-B. (2016). Models of the evolution of fairness in the ultimatum game: A review and classification. *Evolution and Human Behavior, 37*(3), 245–254. https://doi.org/10.1016/j.evolhumbehav.2016.01.001.
Delton, A. W., Krasnow, M. M., Cosmides, L., & Tooby, J. (2011). Evolution of direct reciprocity under uncertainty can explain human generosity in one-shot encounters. *Proceedings of the National Academy of Sciences, 108*(32), 13335–13340.
Feyerabend, P. (1993). *Against method.* Verso.
Figueredo, A. J., Vasquez, G., Brumbach, B. H., & Schneider, S. M. R. (2007). The K-factor, covitality, and personality—A psychometric test of life history theory. *Human Nature, 18*(1), 47–73. https://doi.org/10.1007/BF02820846.
Godfrey Smith, P. (2003). *Theory and reality: An introduction to the philosophy of science.* Chicago: University of Chicago Press.
Gowaty, P. A. (2018). Biological essentialism, gender, true belief, confirmation biases, and skepticism. *APA handbook of the psychology of women: History, theory, and battlegrounds. Vol. 1. APA handbook of the psychology of women: History, theory, and battlegrounds* (pp. 145–164). American Psychological Association. https://doi.org/10.1037/0000059-008.
Gowaty, P. A., Kim, Y.-K., & Anderson, W. W. (2012). No evidence of sexual selection in a repetition of Bateman's classic study of Drosophila melanogaster. *Proceedings of the National Academy of Sciences, 109*(29), 11740–11745.
Griskevicius, V., Tybur, J. M., Delton, A. W., & Robertson, T. E. (2011). The influence of mortality and socioeconomic status on risk and delayed rewards: A life history theory approach. *Journal of Personality and Social Psychology, 100*(6), 1015.
Heisenberg, W. (1971). *Physics and beyond.* New York: Harper.
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. *American Economic Review, 91*(2), 73–78.
Henrich, J., Heine, S. J., & Norenzayan, A. (2010). The weirdest people in the world? *Behavioral and Brain Sciences, 33*(2–3), 61–83.
Heyes, C. (2018). *Cognitive gadgets: The cultural evolution of thinking.* Harvard University Press.
Houle, D., Govindaraju, D. R., & Omholt, S. (2010). Phenomics: The next challenge. *Nature Reviews Genetics, 11*, 855–866.
Humphrey, N. K. (1976). The social function of intellect. *Growing points in ethology* (pp. 303–317). Cambridge University Press.
Ioannidis, J. P. (2005). Why most published research findings are false. *PLoS Medicine, 2*(8), e124.
Kerr, N. L. (1998). HARKing: Hypothesizing after the results are known. *Personality and Social Psychology Review, 2*(3), 196–217.
Klein, R. A., Vianello, M., Hasselman, F., Adams, B. G., Adams, R. B., Jr., Alper, S., …

Bahník, Š. (2018). Many labs 2: Investigating variation in replicability across samples and settings. *Advances in Methods and Practices in Psychological Science, 1*(4), 443–490.

Kokko, H., & Jennions, M. D. (2008). Parental investment, sexual selection and sex ratios. *Journal of Evolutionary Biology, 21*(4), 919–948. https://doi.org/10.1111/j.1420-9101.2008.01540.x.

Kokko, H., Jennions, M. D., & Brooks, R. (2006). Unifying and testing models of sexual selection. *Annual Review of Ecology, Evolution, and Systematics, 37*, 43–66. https://doi.org/10.1146/annurev.ecolsys.37.091305.110259.

Kuhn, T. S. (2012). *The structure of scientific revolutions.* University of Chicago Press.

Leimar, O., & Hammerstein, P. (2001). Evolution of cooperation through indirect reciprocity. *Proceedings of the Royal Society of London. Series B: Biological Sciences, 268*(1468), 745–753.

Loehle, C. (1987). Hypothesis testing in ecology: Psychological aspects and the importance of theory maturation. *The Quarterly Review of Biology, 62*(4), 397–409.

Marlowe, F. W. (2005). Hunter-gatherers and human evolution. *Evolutionary Anthropology: Issues, News, and Reviews: Issues, News, and Reviews, 14*(2), 54–67.

Maxwell, S. E., Lau, M. Y., & Howard, G. S. (2015). Is psychology suffering from a replication crisis? What does "failure to replicate" really mean? *American Psychologist, 70*(6), 487.

Muthukrishna, M., & Henrich, J. (2019). A problem in theory. *Nature Human Behaviour, 3*(3), 221–229.

Nichols, S., & Stich, S. P. (2003). *Mindreading: An integrated account of pretence, self-awareness, and understanding other minds.* Oxford University Press.

Okasha, S. (2006). *Evolution and the levels of selection.* Oxford University Press.

Panchanathan, K., & Boyd, R. (2004). Indirect reciprocity can stabilize cooperation without the second-order free rider problem. *Nature, 432*(7016), 499–502.

Parker, G., Smith, V., & Baker, R. (1972). Origin and evolution of gamete dimorphism and male-female phenomenon. *Journal of Theoretical Biology, 36*(3), 529–553. https://doi.org/10.1016/0022-5193(72)90007-0.

Parker, M. D., Bilder, R. M., Fox, N. J., Freimer, N. B., Kalar, D., Sabb, F. W., ... Chu, I. W. (2008). Cognitive ontologies for neuropsychiatric phenomics research. *Cognitive Neuropsychiatry*.

Pinker, S. (2015). The false allure of group selection. *The handbook of evolutionary psychology* (pp. 1–14). .

Pollet, T. V., & Saxton, T. K. (2019). How diverse are the samples used in the journals "Evolution & Human Behavior" and "Evolutionary Psychology"? *Evolutionary Psychological Science, 5*, 357–368. https://doi.org/10.1007/s40806-019-00192-2.

Price, G. R. (1970). Selection and covariance. *Nature, 227*(5257), 520–521.

Price, G. R. (1972). Extension of covariance selection mathematics. *Annals of Human Genetics, 35*(4), 485–490.

Queller, D. C (1997). Why do femailes care more than males? *Proceedings of the Royal Society of London. Series B: Biological Sciences, 264*(1388), 1555–1557. https://doi.org/10.1098/rspb.1997.0216.

Rand, D. G., Tarnita, C. E., Ohtsuki, H., & Nowak, M. A. (2013). Evolution of fairness in the one-shot anonymous Ultimatum Game. *Proceedings of the National Academy of Sciences, 110*(7), 2581–2586.

Richerson, P. J., & Boyd, R. (2008). *Not by genes alone: How culture transformed human evolution.* University of Chicago press.

Rozenblit, L., & Keil, F. (2002). The misunderstood limits of folk science: An illusion of explanatory depth. *Cognitive Science, 26*(5), 521–562.

Rozin, P. (2001). Social psychology and science: Some lessons from Solomon Asch. *Personality and Social Psychology Review, 5*(1), 2–14.

Saxe, R. (2006). Uniquely human social cognition. *Current Opinion in Neurobiology, 16*, 235–239.

Scelza, B. A., Prall, S. P., Swinford, N., Gopalan, S., Atkinson, E. G., McElreath, R., ... Henn, B. M. (2020). High rate of extrapair paternity in a human population demonstrates diversity in human reproductive strategies. *Science Advances, 6*(8), eaay6195.

Segerstråle, U. C. O. (2000). *Defenders of the truth: The battle for science in the sociology debate and beyond.* Oxford University Press.

Shrout, P. E., & Rodgers, J. L. (2018). Psychology, science, and knowledge construction: Broadening perspectives from the replication crisis. *Annual Review of Psychology, 69*, 487–510.

Simmons, J. P., Nelson, L. D., & Simonsohn, U. (2011). False-positive psychology: Undisclosed flexibility in data collection and analysis allows presenting anything as significant. *Psychological Science, 22*(11), 1359–1366.

Simonsohn, U., Nelson, L. D., & Simmons, J. P. (2014). P-curve: A key to the file-drawer. *Journal of Experimental Psychology: General, 143*(2), 534.

Smaldino, P. E., & McElreath, R. (2016). The natural selection of bad science. *Royal Society Open Science, 3*(9), 160384.

Sng, O., Neuberg, S. L., Varnum, M. E. W., & Kenrick, D. T. (2017). The crowded life is a slow life: Population density and life history strategy. *Journal of Personality and Social Psychology, 112*(5), 736–754. https://doi.org/10.1037/pspi0000086.

Sperber, D., & Wilson, D. (1995). *Relevance: Communication and cognition* (2nd ed.). Blackwell.

Stearns, S. C., & Rodrigues, A. M. (2020). On the use of "life history theory" in evolutionary psychology. *Evolution and Human Behavior* in press.

Tardieu, F., Cabrera-Bosquet, L., Pridmore, T., & Bennett, M. (2017). Plant phenomics, from sensors to knowledge. *Current Biology, 27*(15), R770–R783.

Thornborrow, T., Jucker, J.-L., Boothroyd, L. G., & Tovee, M. J. (2018). Investigating the link between television viewing and men's preferences for female body size and shape in rural Nicaragua. *Evolution and human behavior*. *Vol. 39(5)*. *Evolution and human behavior* (pp. 538–546). Elsevier Science Inc. https://doi.org/10.1016/j.evolhumbehav.2018.05.005.

Tomasello, M. (2009). *The cultural origins of human cognition.* Harvard University Press.

Tooby, J., & Cosmides, L. (1992). The psychological foundations of culture. In J. H. Barkow (Ed.). *The adapted mind* (pp. 19–136). New York: Oxford University Press.

Traulsen, A., & Nowak, M. A. (2006). Evolution of cooperation by multilevel selection. *Proceedings of the National Academy of Sciences, 103*(29), 10952–10955.

Trivers, R. (1972). Parental investment and sexual selection. *Sexual selection & the descent of man, Aldine de Gruyter, New York* (pp. 136–179). .

Wellman, H. M. (1992). *The child's theory of mind.* The MIT Press.

West, S. A., Griffin, A. S., & Gardner, A. (2008). Social semantics: How useful has group selection been? *Journal of Evolutionary Biology, 21*(1), 374–385.

Whiten, A., & Byrne, R. W. (1997). *Machiavellian intelligence II: Extensions and evaluations. Vol. 2.* Cambridge University Press.

Williams, G. C. (2018). *Adaptation and natural selection: A critique of some current evolutionary thought. Vol. 75*. Princeton University Press.

Zietsch, B. P., & Sidari, M. J. (2019). A critique of life history approaches to human trait covariation. *Evolution and Human Behavior* in press.