Evolutionary psychology’s moral implications


Matthew C. Braddock

Abstract In this paper, I critically summarize John Cartwright’s Evolution and Human Behavior and evaluate what he says about certain moral implications of Darwinian views of human behavior. He takes a Darwinism-doesn’t-rock-the-boat approach and argues that Darwinism, even if it is allied with evolutionary psychology, does not give us reason to be worried about the alterability of our behavior, nor does it give us reason to think that we may have to change our ordinary practices and views concerning free-will and moral responsibility. In response, I contend that Darwinism, when it is allied with evolutionary psychology, makes for a more potent cocktail than Cartwright suspects.

Keywords Evolutionary psychology · Sociobiology · Darwinism · Genetic determinism · Free-will · Moral responsibility · Compatibilism

Introduction

In the course of four hundred large double-columned pages, scientist-humanist John Cartwright skillfully maps the interdisciplinary theoretical and empirical space of various evolutionary approaches to human behavior. The theoretical claims, methodologies, and empirical research of Evolutionary Psychology\(^1\) (EP) and

\(^1\) Evolutionary Psychology contends that human beings share a relatively fixed brain architecture consisting of a large number of functionally specialized psychological mechanisms or “modules” (mini-computers) that natural selection designed to solve the adaptive problems faced by our Pleistocene ancestors. These mechanisms process environmental input or “triggers” from a specific domain of experience and respond, among other ways, by contributing to the production of certain behaviors.
“Darwinian anthropology”\(^2\) (DA) are both well mapped and judiciously evaluated, if not always clearly distinguished.\(^3\) Let me review the map in section 2, with concise critical comments sprinkled throughout, and then in sections 3 and 4, I critically evaluate what Cartwright has to say about certain moral implications of Darwinian views of human behavior. I concentrate on the latter topic because Cartwright’s chapter on ethics is brand new to this second edition, and DA’s methodological and epistemological problems have already been ably dissected by other philosophers of biology (e.g. Kitcher 1985; Buller 2005; Richardson 2007), and the topic of evolutionary psychology’s moral implications is not only of special interest to me, but is currently hot and controversial in the philosophical literature.

A critical summary of evolution and human behavior

In Part I, Cartwright begins by sketching a historical narrative of the rise and influence of various evolutionary approaches to human behavior. From Darwin’s arguments for the continuity between human and animal minds to the origins of the idea that the mind is like a computer, to what exactly distinguishes EP from sociobiology, Cartwright sheds light on how we got to where we are today. Since understanding evolutionary approaches to human behavior requires us first to get a handle on evolution, two subsequent chapters familiarize the reader with the basic claims and concepts of neo-Darwinism. Cartwright shows in detail how the shift to a gene-centered view of natural selection, the development of inclusive fitness theory, and work on sexual selection built the theoretical foundations for sociobiological approaches. In an uncharacteristic show of partiality, he throws in his lot with genic selectionism (and gives his reasons), but with characteristic judicious reservation lets us know that the “whole issue is, however, still not resolved.” (p. 43). In the final chapter of Part I, Cartwright crisply characterizes the theoretical claims of EP and DA, outlines and presents problems with their respective methodologies, and bids us be wary of questionable “just-so” stories. His short but sophisticated discussion of the proper testing of adaptive hypotheses is, in my view, one of the best parts of the book.

In two chapters, Part II outlines the evolution of Homo sapiens, critically discusses the origins of human encephalization (brain enlargement), intelligence, and language. He takes an expected ecumenical approach to encephalization and suggests that a good mix of factors, especially social ones, probably brought about

\(^2\) Cartwright groups the methodologically similar approaches of human sociobiology, human behavioral ecology and human ethology under the heading of “Darwinian anthropology” (p. 87). These approaches contend, roughly, that as flexible opportunists we possess domain-general mechanisms that enable us to maximize fitness in our current environments and that we ought to look for adaptation at the level of behavior, rather than ancient psychological mechanisms, by investigating the relation between current behavior and current environmental conditions.

\(^3\) Stephen Downes (2001, p. 587) worries that introductions to evolutionary psychology, like Cartwright’s book (1st edition), tend to mix together research from behavioral ecology and EP, the result of which is confusion and frustration for those who want a clear presentation of the results of these two alternative approaches in order to evaluate their respective merit. This worry legitimately applies to the 2nd edition of Cartwright’s book.
our big brains and our intellectual capacities. Steven Pinker’s arguments for language as an adaptation are reviewed (too) briefly, bullet-point style, and special attention is given to the hypothesis that, due to increasing social complexity, language may have evolved as an adaptive grooming device.

With the debate between rationalism and nativism as backdrop, Part III begins with a useful discussion of the characteristics of Fodorian modularity and EP massive modularity. For novices to EP and those bewildered by the jargon in the literature, the bullet-point presentation of these characteristics will prove useful. A helpful summary of problems with EP massive modularity is also presented at the end of chapter 7. Part III also reviews the evolutionary literature on cognitive illusions and emotions. The main theme is that both phenomena might be understood as adaptive in the environment in which they evolved, though they may not be adaptive in today’s environment.

Part IV covers standard ground on the problem of altruism. Cartwright shows how kin selection, reciprocal altruism, and the capacity for cheater-detection can help explain altruism, while suggesting that game theory offers plausible models for understanding the evolution of cooperation and morality. Part IV also reviews evolutionary approaches to antagonistic and violent behavior. Conflict between parent and offspring, mother and fetus, stepparents and stepchildren, siblings, spouses, and mates are all given space.

To hold up something concrete for inspection and illustration, let me consider Cartwright’s discussion of EP research on infanticide. Drawing on Daly and Wilson’s (1988) innovative work on homicide, he sets out the argument that ancestral conditions may have led to the development of psychological mechanisms underlying infanticide. The idea is that given high infant mortality and low female fertility in the “environment of evolutionary adaptedness” (EEA), certain conditions like that of a woman having a handicapped child or losing the support of the father or family would make caring for an infant extremely fitness taxing, and thus infanticide and its underlying mechanisms may have been selected as a way of enhancing reproductive fitness (p. 218). Though infanticide today may have other (sometimes pathological, sometimes socio-cultural) causes, Daly and Wilson argue that these EEA considerations suggest that originally fitness-enhancing, though often currently maladaptive, psychological mechanisms underlie some infanticide. And Cartwright graphs Daily and Wilson’s statistics showing, for instance, that the frequency of contemporary infanticide decreases as the mother ages, a phenomenon we might expect since infanticide would be less likely to enhance a woman’s reproductive fitness when she has fewer child-bearing years ahead.

But Cartwright does not leave the unwitting reader there to reflect on these established “findings.” He notes that the “results are consistent but do not rule out other interpretations … the fall in risk with age may be a product of coming to terms with the difficulties of child-rearing” (p. 220). I applaud Cartwright’s fairness here and throughout: not out to convert unsuspecting undergraduates, he presents recent empirical research in all its interpretive messiness. Section after section, Cartwright points out the availability of alternative explanations. Such dialectical fairness exemplifies not only epistemic virtue, but also good scientific reasoning, since the latter especially with regard to the psychological and historical sciences typically
consists in the construction and evaluation of alternative explanations. If Evolutionary Psychologists do not seriously consider the best alternative explanations, then they will have trouble convincing others of their own.

Part V looks over evolutionary approaches to human sexuality, with an eye toward identifying typical human mating strategies and showing their prediction from considerations of the asymmetrical reproductive interests of males and females. It is suggested, for instance, that according to work by David Buss and David Schmitt, evolutionary considerations help predict and psychological tests help confirm the fact that men tend to be more philanderous than women. Cartwright also takes stock of the literature on mate choice criteria and provides a masterful overview of the Westermarck Effect that would probably be of interest to moral psychologists and those interested in the evolution of morality.

Part VI introduces evolutionary psychiatry. With characteristic clarity and organization, Cartwright gives us an intellectual toolbox for thinking about mental disorders from an evolutionary perspective (p. 303). Various mental disorders might be viewed, for instance, as naturally occurring defense mechanisms that, though they appear oversensitive, are actually adaptive in the long run. Or they might be viewed as the result of a mismatch between EEA-selected psychological mechanisms and our very different modern environments. He reviews evolutionary research on depression, schizophrenia, autism, and psychopathy.

Part VII begins by sketching five different but not mutually exclusive models of cultural evolution: culture as autonomous from biology, cultural evolution as the natural selection of memes, dual inheritance, gene-culture coevolution, and culture as extended phenotype. Then come a chapter on ethics and an epilogue that wrestle with the moral, socio-political, and broader philosophical implications of an evolutionary approach to human behavior. In the sensible epilogue, Cartwright reviews and refutes the claim that certain unpalatable socio-political implications—e.g. eugenics, racism, sexism—follow from Darwinism. He also argues that Darwinism is not incompatible with nor in tension with left-leaning liberal political philosophy and that the nurture-only “Standard Social Science Model” (SSSM) faces more unpalatable socio-political implications than might initially be thought. Cartwright presents sound arguments here: Darwinism is not as philosophically nefarious as many uncharitable critics suggest. But in the ethics chapter, he goes further and suggests, in step with Janet Richards (2000), that apart from making dualism and its associated libertarianism implausible, Darwinism is not very philosophically powerful. In sections 3 and 4, I try to show that Darwinism in combination with Evolutionary Psychology makes for a more potent cocktail than Cartwright thinks.

**EP and the alterability of behavior**

Cartwright confronts an enduring objection articulated by the late Stephen J. Gould, call it the Objection from Genetic Determinism: “If we are programmed to be what we are, then these traits are ineluctable. We may, at best, channel them, but we cannot change them either by will, education, or culture” (p. 345). Originally the
criticism (fairly or unfairly) was put to sociobiologists, but in recent decades we have witnessed its unsurprising extension to EP. This objection, and presumably those like it, can be “easily dismissed” according to Cartwright because genetic determinism is obviously false and because “Evolutionary psychology tells us about the functional origin of emotions, drives and decision-making algorithms but not how compelling any of them are” (p. 345).

Cartwright’s response (fairly or unfairly) makes Gould’s Objection from Genetic Determinism look silly: everyone knows nature and nurture both contribute to behavior. Cartwright is right here and elsewhere in the book to press the point that EP in no way entails genetic determinism. That the environment matters to EP is evinced by its fundamental claim that the psychological mechanisms underpinning human behavior are adaptations to our ancestral Pleistocene environment. For to claim that such mechanisms are adaptations to ancestral conditions is not to say they are impervious to environmental influences. Novel circumstances could possibly change the character of these EP mechanisms, so that the usual triggers that activate them might produce very different outputs. New circumstances and selective pressures might even eliminate these mechanisms. And furthermore, given our stellar capacity to construct our own niches, we might be able to socially engineer our circumstances so as to remodel these mechanisms ourselves. All such possibilities are consistent with modern evolutionary theory.

And though these mechanisms are alterable, even if they were not, Evolutionary Psychologists would still not be committed to the claim that their behavioral output is fixed, because we would still have some measure of control over the environmental triggers that activate the mechanisms. For instance, we might eliminate the triggers that activate mechanisms underlying violent behavior or we might proliferate the environmental cues that trigger cooperative behavior. Thus, there appears to be no good argument to the worrisome conclusion that EP entails the ineluctability of our psychological traits or behavior.

Yet can we dismiss Gould’s objection in a few sentences and be done with concerns about alterability? Cartwright and others in my view, are too quick here. I suspect that aside from unjustifiable worries about genetic determinism, there are serious concerns about practical alterability that must be filled out and addressed. Though the psychological mechanism and their behavioral outputs are alterable in principle, they may nevertheless be very difficult to alter in practice. And if (a hypothetical I will soon argue is actual) EP implies that the mechanisms underlying (say) infanticide and philandering, as well as the behaviors themselves, are practically unalterable, such an implication would be cause for worry. Why worrisome? Because if it is very difficult to eliminate these sorts of behaviors, then many will not unreasonably be inclined to think we should not waste our collective resources trying to do so. The implication, on this argument, would be that if policymakers accept EP claims and findings about (say) the origins of infanticide or philandering and endeavor to let EP (as good science) inform their policymaking, then though they would of course attempt to protect society from the effects of such behavior, they would probably do little to nothing in the way of attempting to

---

4 See sections I and II of Lewens (2003).
modify the behavior itself. If EP were true, then such conservatism about the practically unalterable behavior underpinned by EP-mechanisms would be the most reasonable course of action. Yet if EP is false, then were it to be accepted and were it to influence policymaking in a conservative (do nothing) direction, the potential human cost could be significant. Social resources that could have been allocated toward (say) minimizing infanticide might instead be thrown at some lesser cause.

But for this broad argument to be plausible, the antecedent of the conditional outlined in the previous paragraph must be defended. Why think that EP-postulated mechanisms and their associated behavioral responses are in fact practically difficult to alter? The argument for this might be sketched as follows:

If

(1) EP-postulated mechanisms are causally important in the production of behavior,
(2) it is very practically difficult to alter the character of these mechanisms,
(3) and it is very practically difficult to eliminate the environmental triggers that activate these mechanisms,

then,

(4) disturbingly, it would be very practically difficult to eliminate the behavioral responses these mechanisms produce.

Let me briefly comment on each premise.

Cartwright evades Gould’s objection by pointing out the obvious fact of gene-environment interaction. But he goes further and, as I see it, tries to inoculate EP against unpalatable practical implications by saying, with Janet Richards (2000, Ch. 5), that EP “tells us about the functional origin of emotions, drives and decision-making algorithms but not how compelling any of them are” (p. 345). The suggestion is, contrary to my premise (1), that since EP does not make claims about the causal strength of EP-mechanisms in the production of behavior, then we should set aside any worries we might have about the implied alterability of these mechanisms and their associated behaviors. After all, if these mechanisms do not affect our behavior much, then we can swallow their resistance to alteration with ease.

But the problem with Cartwright’s apparent rejection of premise (1) is that Evolutionary Psychologists embrace it and herald this discovery as the foundation for a new revolutionary paradigm in psychology. They are not agnostic about the causal importance of these mechanisms, nor do they present them as negligible or relatively significant causal influences on our behavior; rather, they think they have tapped into the well from which human behavior typically springs and are vocally trying to disillusion the SSSM-believing social sciences and broader public. My broad conditional argument is that if we accept EP’s claims and findings, then worrisome socio-political implications of the kind outlined above may reasonably follow. And premise (1) is a foundational pillar of EP’s research program, so its rejection cannot be the way to shield EP from the argument from practical unalterability.

Why think premise (2) is correct? Given the nature of EP-mechanisms qua adaptations, it is at least biologically possible that novel circumstances might
modify the character of these mechanisms. Nevertheless, given that we do not know which circumstances would significantly affect the character of such mechanisms, it then becomes in an important sense practically difficult to alter them. Consider an illustration. It is theoretically possible to produce and harness cold fusion with the resources and technology currently available, but it is practically difficult to do this because we do not (as of yet) know how to do it. Hence, unsurprisingly, cold fusion science is treated as “fringe” and policymakers are not quick to invest social resources in its pursuit. In the same way, it is theoretically possible to change the character of many EP mechanisms easily with the resources and technology available, but it is practically difficult because we do not (as of yet) know how to change them. Thus, if policymakers accept EP claims and findings, they will probably not be inclined to devote social resources toward the reform of EP mechanisms.

“Even if it is difficult to alter the character of EP mechanisms,” one might respond, “the undesirable behavior (e.g. infanticide, philandering) might nevertheless be minimized with practical ease by identifying and systematically eliminating the environmental triggers that activate these mechanisms.” Premise (3) is my counter to this proposal. The environmental triggers that activate these mechanisms cannot easily be eliminated, because (i) the many triggers cannot (as of yet) be easily identified and, even if they could be, (ii) it is hard to see how policymakers could easily eliminate the bulk of such triggering situations without resorting to extremely controversial political measures like a radical redistribution of wealth (e.g. so that all pregnant women in situations that may trigger infanticide-underpinning mechanisms are provided with the long-term resources necessary for caring for their children) or undue interference with basic human rights and freedoms (e.g. so that men do not face situations that trigger mechanisms for philandering behavior). Such an undertaking would not be easy, given our current socio-political circumstances, and thus would likely not be pursued by today’s policymakers.

Let me summarize: since it would be very difficult to alter the character of EP-mechanisms and to eliminate their associated triggers, it would thus be difficult to eliminate the undesirable behavioral responses that these mechanisms have a causally significant role in producing. The worrisome practical implication is that, if policymakers accept EP claims and findings, they will probably not endeavor to alter such behavior. And yet Cartwright, as well as many others, point out the falsity of genetic determinism, highlight the fact that the environment plays a part, and think they are done with Gouldian-style objections. If they carry on like this, these concerns about practical alterability and its likely socio-political implications will probably stick around, and not without reason.

**EP and free-will/moral responsibility**

Cartwright confronts the worry that neo-Darwinism (ND) undermines free-will and moral responsibility (pp. 345–348) and argues, in the vein of Janet Richards (2000, Ch. 6), that ND sits comfortably with all that we cherish and can legitimately want
with regard to these two related subjects. He appreciates the strength and widespread distribution of libertarian intuitions, but is impressed by the old argument that an appeal to indeterminism—that ontological appeal to uncaused events—will not solve the free-will problem for the simple reason that nobody can be intelligibly responsible for uncaused events. What then becomes of moral responsibility in a neo-Darwinian world? Here is how Cartwright spells out the implications:

If free-will has thereby been demonstrated to be incompatible with determinism and indeterminism [due to libertarian intuitions and the old argument described above concerning indeterminism], it seems that we have ruled it out of existence altogether or that we have an imperfect conception of what ‘free will’ actually means. … [I]t is worth considering Richards’ suggestion that we need to distinguish between two types of responsibility. One is ‘ordinary’ responsibility, which means being of sound mind and not under the influence of drugs or disturbed in some way; the other is ‘ultimate’ responsibility, which is something to do with a chain of causation. As our understanding of genetics, development and enculturation becomes more sophisticated, so we will achieve a better understanding of the forces that make us what we are. It does not follow, however, that praise and blame for moral and immoral acts respectively become unnecessary. Instead, they both serve as essential ingredients to inject into the decision-making of all of us. They become ‘instrumentally necessary for the maintenance of social order’ (Richards 2000, p. 151). (p. 348)

It is clear that he thinks moral responsibility can be retained, but only in its compatibilist or instrumentalist forms; “ultimate” responsibility of the libertarian stripe is surely out. With some reconstruction on my part, his reasoning might be outlined as follows: our ordinary responsibility practices and attitudes are perfectly compatible with determinism, but even if we have libertarian intuitions and think they are not, all is not lost. For libertarians would still have good pragmatic reasons to retain our ordinary responsibility practices and attitudes, namely because the latter would be instrumentally valuable in maintaining social order, and we all want social order, right? So even if we do not find compatibilism acceptable, there is still no good reason to think ND has deeply revisionary and worrisome consequences for the way we think about moral responsibility.

In response, first, I take it for granted that an instrumentalist view of moral responsibility is deeply revisionary and worrisome. If this is where the libertarian is driven and if libertarian intuitions are strong and widespread, as Cartwright suggests, ND has strongly revisionary implications the magnitude of which Cartwright does not seem to appreciate. Second, he thinks ND sits comfortably with standard compatibilist views and our ordinary responsibility practices, but I suggest that ND in combination with EP has important revisionary implications that he does not suspect.

Let me draw out one implication by beginning with his observation that compatibilists and most people accept various intuitive distinctions between the responsibility of ordinary rational deliberators and that of other sorts of people (e.g.
kleptomaniacs, those under the influence of drugs or alcohol) whose deliberations are impaired or strongly influenced by various internal and external constraints. The intuitive idea is that the latter group’s behavior is not fully responsible behavior because of the way in which it is caused. Yet it seems to me, and here is the crucial premise, that the behavior characterized by Evolutionary Psychologists as being underpinned by various psychological mechanisms is relevantly similar in the manner it is causally produced to the behavior of those who are paradigmatically not as responsible for their actions as the rest of us ordinary rational deliberators are for ours. Let me put it this way: on the continuum of morally responsible behavior—from the behavior of paradigmatically non-responsible agents like kleptomaniacs at one extreme to that of ordinary rational deliberators at the other extreme—behavior produced by EP mechanisms lies more toward the kleptomaniac extreme than we may have previously thought (i.e. before EP findings), though how close to that extreme depends on how causally important the mechanism was in the production of a specified behavior. Given that Evolutionary Psychologists typically ascribe significant causal strength to these mechanisms in the production of behavior (my premise (1) in section 3), in order to be consistent with the intuitive compatibilist distinctions we make, we should consider the operation of such mechanisms as a mitigating factor in personal and collective decisions concerning moral responsibility. ND, when combined with EP, threatens to significantly alter the degree of compatibilist freedom and moral responsibility we ordinarily assign certain people.

Consider a specific case. Before EP research on infanticide, we may have thought that a mother who killed her newborn infant was (barring postpartum depression and other traditional mitigating factors) just as responsible as any ordinary rational deliberator for the death of her newborn. Yet if Daly and Wilson and others are right in suggesting that infanticide may sometimes be in large part the behavioral result of environmental triggers acting on an innate evolved psychological mechanism designed by natural selection to enhance her ancestors’ reproductive fitness, the mother does not seem as responsible as she would have been were that information not available. Just as potent drugs may function as strong internal influences on behavior, so these psychological mechanisms may strongly influence behavior. And if an individual’s responsibility is mitigated, though not necessarily negated, in the former case, then why not also in the latter case? What principled distinction is there between the sorts of influences (drugs, alcohol, mental impairment) recognized intuitively and formally in the criminal justice system as factors mitigating responsibility and the sorts of influences identified by Evolutionary Psychologists? At this point, I do not think such a distinction can be made.

To add up: from a consideration of intuitive compatibilist distinctions that are reflected in our ordinary personal and socio-legal practices, we may reasonably infer that we ought to count behavior underpinned by EP mechanisms as satisfying conditions for diminished responsibility and that the diminution should be significant since, as Evolutionary Psychologists contend, these mechanisms are causally significant influences on behavior. If EP were accepted, these reasonable inferences would probably bring about deep revision in our ordinary and socio-legal views and practices with regard to the assessment of people whose behavior is produced in this way. Were EP to be accepted, this implication for responsibility...
drawn, and if EP were to turn out to be false, the consequences could be disturbing: by making use of an “EP legal defense,” many might get off “easy” in criminal cases while their relevantly similar counterparts, who could not appeal to such a defense, might be (justly) locked up for a long time. This potential injustice should be taken seriously.

Notwithstanding Cartwright’s Darwinism-doesn’t-rock-the-boat approach, EP does have consequential implications and were EP to be accepted and yet turn out to be false, those implications may have significant human costs. Like the argument from practical unalterability in the previous section, this argument from intuitive compatibilist distinctions gives us reason to raise the standards of evidence that EP must meet in order to be accepted and to be allowed to inform public policymaking, an outcome that exacerbates EP’s already existing methodological and epistemological problems.

Conclusion

Many thanks to Cartwright for publishing this unique interdisciplinary, fair-minded, and eminently readable introduction and, also, for boldly taking up the question of the moral implications of Darwinism. But on the latter topic, much more needs to be said.

References

Richardson R (2007) Evolutionary psychology as maladapted psychology. MIT Press, Cambridge, MA