Ethics without Free Will

Aaron L Bramson

Abstract

Moral theories typically rest upon the assumption that conscious deliberation plays a causal role in action; however, a growing body of scientific evidence supports a physicalist account of causation that leaves no causal role for mental activity. In response, I develop a moral theory (including moral truth, motivation, and meaning) that excludes considerations of conscious free will. I then consider how an acceleration in social evolution compared with biological evolution affected our moral knowledge and motivations. The idea of a normative model is offered to replace normative ethical theorizing, and finally various problems of the theory are identified for future work.

1 Introduction

The combination of 1) appealing reductionist arguments from metaphysics and 2) recent scientific discoveries in psychology, neurology, bio-chemistry, and physics gives strong support to a physicalist account of behavior devoid of any ghost in the machine. If the thesis that mental states are irrelevant to human behavior is defendable, then any philosophical ethical theory that hinges on humans' ability to act intentionally fails to provide a proper foundation for moral truth, meaning, and motivation. Much of the recent work in ethical theory has focused on the relation of reasons to motivation and action, but only in the context of particular (though varying) models of psychology. By taking seriously the evolutionary origins of emotions, ethical intuitions, and physical responses I develop a theory of moral meaning, truth, and motivation independently of humans' ability to act through conscious free will.

The question of the evolution of human society and morality is an old one. When and how did humans develop the ability to regulate our activity according to social norms? Are humans unique in our development of morality and ethical principles? What facts about human nature (the biologically determined aspects of human beings) are relevant to moral thinking and action? This paper is an attempt to use principles from modern evolutionary theory to flesh out something like the following story of Machiavelli's.

> These various kinds of government came into existence among men by chance, for in the beginning of the world, the inhabitants being few, they lived dispersed for a time in the manner of beasts. Then, as the population increased, they drew together and, the better to defend themselves, they sought out the strongest and bravest one among them, made him their leader, and obeyed him. From this

> beginning came a recognition of what is proper and good, as opposed to what is pernicious and wicked. Seeing a benefactor injured, men came to feel hatred and sympathy. They censured the ungrateful and praised those who showed gratitude. From this came the recognition of justice.[12, p.92]

In what follows, the phrase 'free will' shall be used synonymously with 'conscious decision making', 'deliberation', 'intentional action', 'choice', and any other phrase meaning that mental states or processes affect human behavior, acts, or physical states. While some authors have created distinctions among these terms [15], they all imply a role for psychology which is irrelevant to the theory presented here. Also, I will slip between the use of 'physicalism' and 'reductionism' to refer to a physicalist reductionist view, i.e. I will use one or the other term to refer to the joint theory except where noted.

The paper is organized in the following way. I first present some brief arguments and evidence for the validity of the claim that human psychology is irrelevant to human action. I then outline how and why biological evolution equipped us with the moral feelings that we have and how one might construct meta-ethical and philosophical moral theory from such an explanation. The effects of accelerating social evolution on morality are then analyzed in accordance with this biologically based moral framework. Normative ethics is recast as being analogous to psychology in supplying models for explaining and guiding action without the need for causal linkage. And finally I consider what has been left out of the moral theory presented here and what roles society and free will are needed to fill.

2 Some Arguments Against Free Will

Since the main thrust of the paper is to determine /textithow far one can get in constructing a moral theory that does not include intentional or deliberative action as a component, even if the premise that humans do not exhibit free will cannot be sufficiently defended, the primary thesis remains unscathed. Through the analysis below, we will see the role that we need intentional action to play in a moral theory, i.e. what gaps a moral theory without free will has and how reason may fill these gaps. However, to help motivate developing a physicalist moral framework, I will provide brief summaries of three arguments against human conscious reasoning's efficacy to action.

2.1 Argument from Supervenience

Donald Davidson introduced the idea that "mental characteristics are in some sense dependent, or supervenient, on physical characteristics. Such supervenience might be taken to mean that an object cannot alter in some mental respect without altering in some physical respect. [5]" Accepting such a thesis *prima facie* leans heavily towards a reductionist (physicalist) theory of the mind[1]. I wish to deploy the following corollary of supervenience due to Jaegwon Kim.

S: the supervenience of a class of properties M upon another class B actually *entails* that M is reducible to B [11].

I refer readers to the original paper for the proof of this claim. Once accepted, holding that the mind supervenes on the brain commits one to mental states being reducible to brain

[1]Davidson himself did not accept such a conclusion; however, further discussion of the adequacy or inadequacy of supervenience for establishing physicalism would take us too far a field.

states.

The possibility of free will requires that an individual's mental processes can cause changes in an individual's physical states (let's say, in the individual's brain). Let M be a person's mental state and B be the person's brain state at time t. Let M^* and B^* be the person's mental and brains states, respectively, at time $t+1$. I take the free will hypothesis to be the following claim[2]:

FW: the change from M to M^* *caused* the change from B to B^*.

But on a physicalist account, natural laws of bio-chemistry and physics are sufficient for explaining the brain state change from B to B^*.[3] So taking together i) supervenience of the mind on the brain and ii) the sufficiency of a physical explanation for changes in brain states we must reject **FW**.

2.2 Argument from Psychology

If we are interested in the role of mental activity, a natural place to look is the field of psychology. Wegner's book *The Illusion of Conscious Will*[16] provides a historical overview of psychology's investigation of conscious will with the modern perspective that it does play a causal role in our behavior.[4] Multiple examples from his text speak against a theory of intentional action of the kind deployed by moral philosophers.

One example from the hypnosis literature recasts an experiment by Albert Moll which demonstrates that an individual "embarking on a behavior for which no ready explanation came to mind, freely invented one in the form of a prior intention. [16, p.149]" This faculty is precisely what Michael Smith calls upon to provide us with normative reasons for action [15, p.132]. Other examples demonstrate that even when there exist sufficient explanations of behavior independent of consciousness, the authorship emotion[16, p.325] kicks in and ascribes the action to our mental causation. There is likely a good evolutionary explanation for our having such a reaction (Wegner provides one), but discussing it will be left for future work.

2.3 Argument from Neuro-Biology

There is a great deal of anecdotal evidence from sports of, for example, runners starting a race before consciously hearing the starter pistol's firing [4]. There is evidence from brain damage cases where people perform a very complicated series of actions that normally elicits sensations of decision making, but fails to do so in the brain-damaged individuals[16]. No matter how much mental activity we experience, it is becoming increasingly clear that it just the unconscious activity of the brain that causes the body's motion. There seems to

[2]There are other formulations of what free will amounts to, some quite a bit weaker than the claim FW. I do not wish to set up a straw man nor trivialize these other accounts; FW captures the claim implicit in moral philosophy that this paper aims to show is irrelevant to moral truth and motivation.

[3]While it is true that due to the complicated mechanisms involved in brain activity no such full account is practically possible in the foreseeable future, in light of the success of physical bio-chemistry and the enterprise of scientific explanation in general the burden of proof for the inadequacy of such an explanation rest squarely on the non-reductionist theories.

[4]Wegner does not claim that mental events fail to be causal, but only that consciousness is. His accounts are therefore compatible with the argument from supervenience against FW above (if all mental events are indeed reducible to physical events), but they are not a mere restatement of it.

be no room for conscious free will in the determination of our actions.

This does not mean that the mental sensations of deliberating and deciding are uncorrelated to our actions. Such a position would be "tantamount to supposing that the connection between what we decide to do, on the basis of rational deliberation, and what we do do, is altogether contingent and fortuitous. And that is patently absurd. [15, p.132]" Indeed the connection seems to be a very close one; mounting evidence shows that the connection is that what we actually do (or did) determines what we decide on the basis rational deliberation. The evolutionary explanation for our having a consciousness also needs to provide an explanation of why it is tied so closely to our actual behavior instead of constantly daydreaming.[5]

2.4 The Persistent Illusion of Conscious Will

Kant said that even though he could not prove the existence of free will, we cannot deny the sensation of conscious deliberation, and the seeming efficacy of decision making on our actions[10]. One might expect an evolutionary account of ethics to include a story as to why the moral sense (or consciousness at all) has evolved[6]. But Kant also held that our rational free will is a necessary condition for establishing any moral theory[10]. The current work assumes the opposite; one can construct a theory of moral meaning, truth, and motivation wherein conscious free will is irrelevant (even if present).

3 Evolution of Emotions Fit for Morality

I cannot expect the resolve the debate over the role of evolution in the formation of our intuitions (or emotions or feelings) and which impulses trigger them in this short paper. The discussion spans multiple disciplines and an even greater number of conclusions. A summary of the evidence for and against my view of the evolution of humans and society will not be attempted. The argument I wish to make here shall hang minimally on any particular story of evolutionary history and more on necessary properties of systems wrought from an evolutionary mechanism. One distinction particularly relevant to the account given below is a hypothesized schism in social versus biological evolutionary timescales and its role on the evocation of morally relevant sentiments and emotions.[7] I propose that even though social situations absent when behavior was determined solely through biological evolution elicit (seemingly appropriate) emotive responses, the mechanisms that produce emotive reactions evolved in response to selection pressures no faster than the biological timescale. The first step in developing of a moral theory based on evolution, therefore, starts when biological evolution and social evolution (henceforth 'bio-evolution' and 'socio-evolution') progressed on the same timescale.

[5] There is also an argument from meme theory that liberates our consciousness from the duty of deciding our actions (or other thoughts). The idea, originated by Richard Dawkins[6] and expanded by (among others) Susan Blackmore [1], has our actions determined by self-replicating causal entities (called 'memes') that spread via the behavior (including speech acts) that they induce in us carriers.

[6] A recent work that does address some of these questions is Daniel Dennett's *Freedom Evolves*[7], though his stance is quite askew from my own view regarding the role of free will in action.

[7] To claim that some emotions are not relevant requires some prior moral theory to determine what is relevant. The evolutionary approach taken here has that all emotions are morally relevant. There is, however, a distinction between morally *relevant* feelings and moral feelings. The former are any emotions people have that must be considered in our description of how individuals ought to act and why. The latter are emotive responses to violations of and exemplars to how individuals ought to act. Modern preference-satisfaction act-utilitarianism is an example of a cross-level moral theory that includes the full manifold of emotions in determining what is ethical, insofar as emotion is relevant to utility.

3.1 The Original Position

Humans, like many animals, are social creatures that evolved to live in groups with specialized roles for different members of the group. Some social mammals, e.g. wolves seals, lions etc., have a social structure that is instinctive (i.e. hard-wired in their biology). Development, nutrition, accidents, and other such factors certainly play a role in determining the particulars of any animal social structure, but the underlying structure of animal societies in genetically determined and changes in it must occur at the bio-evolutionary timescale. Evolutionary theory tells us that (say) wolves have such a social system because wolves that acted in accordance with that system produced more childbearing offspring than wolves that acted otherwise.

In the history of human (biological) evolution there was a time when changes in our social structure were also bound to the bio-evolutionary timescale (I will henceforth refer to human ancestors existing in this time period as 'protohumans'). At that time, emotions evolved as reliable guides to the actions that tended to improve evolutionary fitness. Had these emotions been unreliable guides, that is to say if emotions motivated acts that frequently prevented the actor from performing well enough to rear reproductive offspring, then the genetic material that encoded the emotions would not have been passed on. Some unfit responses may have "piggy-backed" on successful traits, but as long as the selection pressure is strong enough and the timescale is long enough, all but a few minor deleterious characteristics will have been removed. What protohumans were left with was a coherent meshwork of biologically determined social structures and emotive responses appropriate for propagating those structures insofar as maintaining those structures benefited the genetic fitness of the individuals doing so.

3.2 Biological Foundations of Morality

Given a social animal with a biologically determined social structure it seems unproblematic to ascribe a normative system identifying what acts, feelings, institutions, etc. are condoned, rewarded or reproved (see Moral Models and Normative Ethics below). It is less obvious that a meta-ethical theory and a philosophical moral theory for moral truth, motivation and meaning can be so constructed. To start we need to recognize that the moral intuitions, feelings, physical responses, and social institutions (norms) of protohumans all coevolved with each other, the environment, and all the other aspects of protohuman existence. Thus whatever emotion was felt, that the emotion persisted in the population as a response to whatever elicited it implies that the emotion felt was the appropriate (in terms of reproductive fitness) emotion for the situation that elicited it.[8]

The cohesive web of social structure, moral feelings, reflexive responses, food gathering techniques, mating rituals, number of toes, eye colors, and everything else about a species (including its interaction with other species) has coadapted to each other and the environment form a complex system of interconnected relationships. Individuals, their behaviors, motivations, etc. may hold moral value, but it is at the system level that moral properties must be evaluated because that is the level at which fitness is meaningful.[9] A behavior is

[8] This claim glazes over the details of mutation and other forms of genetic experimentation that allows for evolution in the first place. Some emotion may be *in*appropriate for a given system, yet be fitness increasing. So instead of being weeded out of the population, the once inappropriate feeling will eventually become more prominent in the population and then become the standard of appropriateness. The processes is continually changing what attitudes are fit to what acts, albeit at the excruciatingly slow pace of biological evolution. (More in Conclusions below.)

[9] Moral properties therefore have a similar status as truth properties in a coherentist epistemic theory[2,

selected for through pressure exerted by the whole arrangement of other elements of the system, even though it is the individuals' fitness that is being measured. At any give point in time, the behaviors that agents enact are precisely the ones that yielded the highest fitness in the evolutionary history of that system.

If we accept that the correct use of the moral 'ought' implies 'can' then because, *ex hypothesi*, protohumans' behavior (like present-day wolves' behavior) is fully determined by their biological make-up we are also forced to accept that either 1) what they do do is somehow what they morally ought to do or 2) that the moral 'ought' simply fails to apply to deterministic creatures. The second route is Kant's famous dictum, which, even if not at the core of their theories, seems to have been accepted by every moral philosopher before or since. I now examine what sort of philosophical moral theory one can construct by accepting (1).

3.2.1 Moral Motivation

Evolutionary theory provides a clear explanation of why one has certain motivations to act in certain circumstances: one's progenitors had the motivation in those circumstances while others did not and those with the motivation became more prevalent via selection pressure for (among other things) having that motivation. In this sense we can say that selection pressure provides *reasons* for having certain motivations, even though evolution is not a teleological process. And biologically adapted behavior provides reasons^N to act in the corresponding way (following the Nagelian [13, p.331-332] account of reasons). Searching for and consuming food is an adapted response to requiring nourishment. The physical feeling of hunger and the impetus it provides for action are adaptations that benefited those creatures in which it evolved. Animals have all and only those motivations to act that tended to improve their reproductive fitness.[10] They have reasons^N to eat when they are hungry, and these reasons^N are necessarily motivating; but the necessity comes from the evolutionary explanation for them having the motivation in the first place. The *reasons* for having the motivation are precisely the factors of the system that produce selection pressure on individual fitness.

The *reasons* and reasons^N for feelings of guilt, jealousy, comradery, etc. are the same as those for feelings of hunger. Such feelings (impulses, emotions, reactions, etc.) motivated and hence produced behavior that was selected for in the coevolving system. If we accept that a social system of protohumans is capable of exhibiting morally evaluable behavior, then that behavior is motivated by reasons^N for which individuals have *reasons* to be motivated by. Furthermore, these internal and external reasons are necessarily bonded to one another.

p.250]. This is opposed to being emergent properties of the system; properties of the whole itself by virtue of mereological properties.

[10] I will reiterate the caveat that some motivations to act may not have, in fact, improved fitness but rather "piggy-backed" along with other adaptations that did. Evolution is not a truly maximizing process, so detrimental behaviors may persists as long as they are not too destructive to the systems. Biological history and computer simulations have also shown that mutations that are deleterious at their onset sometimes become advantageous when the evolutionary environment changes. Furthermore, minor deleterious changes are sometimes necessary genetic gateways to much larger beneficial adaptations. Also note that this only holds true, insofar as it does hold true, for motivations and other traits that actually did evolve.

3.2.2 Moral Truth

Statements about the appropriateness of any protohuman social structure or individuals' reaction within that structure are true if and only if the protohumans have that social structure or individuals exhibit that reaction. One can assign a truth-value to a moral claim about an individual act or emotion, but its truth or falsity hinges upon its matching the features required of that component of the coevolved system. Hence the truth or falsity of a moral claim can only evaluated in terms of role the evaluated element plays in the cohesive web of interrelationships of the system. The moral epistemology naturally implicated by the coevolutionary framework is a coherentist epistemology[2, p.250].[11]

An example will help demonstrate how moral truth-values are assigned under the current theory. Some individual in the protohuman population, call him Bob, is jealous of another's good fortune in the afternoon hunt. Two typical morally relevant questions we can ask of Bob's response are: "Is his jealousy justified?" and "Ought he feel jealous about this situation?" A theory of moral truth needs to tell us whether the answers to these questions are true or false.

Because jealousy is Bob's *de facto* response to the situation, and that response evolved along with the social system, it is clear that his jealousy *is* justified in that situation. That kind of situation is precisely the situation in which feeling jealous is justified. The feeling of jealousy evolved to elicit the appropriate response in individuals in precisely the situations in which protohumans do, in fact, feel jealous. That Bob feels jealous in that situation is therefore sufficient to assert "Bob is justified in feeling jealous" is true.[12]

The question whether he ought to feel jealous in that situation is answered in a similar manner. Feeling jealous will produce the behaviors that tend to increase the individual's fitness in the system he finds himself, whatever behavior that happens to be. Since the emotive response and behavior it elicits are linked via the evolutionary process, behaving otherwise will harm his fitness and fail to contribute his part to the system that evolved to have jealousy as the appropriate response to that situation. "He ought to feel jealous in that situation" is therefore also true. To say that he ought not feel jealous would be to claim that the system would somehow benefit from a different feeling-behavior pair; but if a different behavior were better within that system, then evolutionary forces would have made this other behavior the elicited one. (Problems with the backwards looking approach taken here is examined further in the Conclusions section below.)

3.2.3 Moral Meaning

Our concept of morality is what it is because (biological and social) evolution favored individuals with such a conception. The meaning of moral terms is grounded in our moral intuitions, which in turn is grounded in our biology to some extent. But even if our concept of (say) *the good* formed via the same evolutionary process as our concept of *acorn*, this does not imply that the meaning of 'good' can be cached out in a similar way as the meaning of 'acorn'.

The current theory of ascribing a positive moral value to the behavior of agents within

[11] As noted in a previous footnote, all aspects of the moral theory presented here fall along the same lines as a coherentist epistemology because all moral properties (motivation, truth, and meaning) apply to individual elements of the system but only by virtue of their role in the whole system.

[12] This assertion assumes, of course, that Bob is a typical member of that protohuman society.

a system just in case the behavior evolved as appropriate (in terms of biological fitness) for the system blurs (or perhaps equates) the descriptive and evaluative meanings of moral terms identified by Hare [9]. It is also at risk of embodying the naturalist fallacy insofar as what is morally right is cached out in terms of what is the biologically evolved response.

My best response to these problems at present is to recommend that we examine what concepts of morality the protohumans could have and whether the meanings of moral terms that they would have (if they had language) would match the moral value-giving properties that I have defined above. It would be unfair to impose *our* conception of morality on the protohuman social system just as it would be to do so on a pack of wolves. I want to know what can be learned about developing a moral foundation by taking seriously the question, "What concept of morality would a wolf have if wolves had a concept of morality?"

That it would be different from our own is quite clear. Wolf behavior can be, in many cases at least, easily associated with expressing a pro or con attitude (growls, whimpers, grooming, etc.). Wolves would certainly have a concept of morality that falls in line with their social system and the interconnected relations they hold to their environment and each other. While modern humans may find a social system based upon an alpha male hierarchy morally reprehensible, arguable a wolf would not think so. A wolf would have to have a concept of morality consistent with providing a small group of the pack a larger share of the resources and punishing defectors in the exclusive mating rights of the alpha male.

I can image that the lowly males in the society feel repressed by the hierarchal system; that they feel that they are not getting their fair share. But that is certainly just imposing my modern human moral concept onto a system for which it was not intended. The moral concepts that a wolf would have are precisely the ones that would value the behaviors that do, in fact, elicit pro attitudes and devalue those that elicit con attitudes in actual wolf behavior. And the pro and con attitudes evolved to be appropriate for their social systems according to the story already presented in the above sections.

Would there be a shared moral concept between protohumans and wolves if they could have and discuss such a concept? Though the criteria for morality are different in the sense that different actions will elicit pro and con attitude-related responses in the two populations, both populations could accept that 'morally justifed' means eliciting behavior that has evolved to provide the highest fitness for their respective societies. The similarities and differences are analogous to those in Hare's missionaries and cannibals example[9]. Both societies value what they do for the same *reasons*, it is just that differences in their biological make-up and in their environments have provided them with different particular reasons^N.

3.3 But Is It Morality?

Perhaps you are not convinced that the above account actually amounts to a philosophical and meta-theory of morality. What I have done is present explanations for why individuals that evolved within a system have the responses to situations that they do and why those responses are the best responses to have for those individuals. I have attempted to convince you that what such individuals would find morally good (if they were capable of such a conception) is the behavior that evolution would favor as being considered morally good. The coevolutionary tale given is not a "just so story", it is a "if it could have been otherwise then it would have been" story. Evolution only takes one path to get to where it ends up. The preceding account does not depend on any notion or faculty of free will to provide the

grounding of morality. But it is still not clear whether a causal account of the origins of moral motivation, a fitness account of moral truth, and a hypothetical account of moral meaning can form the basis of a moral theory. Even less clear is what needs to be added or changed to make it one. If my account can amount to a genuine philosophical theory, then it is a very different one from the ones with which I am familiar, thus making its status difficult to evaluate.

4 Persistence of Emotions Unfit for Society

I now turn to the question of what insight establishing such a moral theory for protohumans yields for modern human morality. At a certain point in the evolution of modern man the speed of social evolution outpaced our biological evolution. Biological evolution continued (and still continues) to exert selection pressure on human adaptation, but genetic make-up could not change quickly enough to adapt our biologically grounded response mechanisms to the environment we were creating for ourselves. The slide from the period of lock-step bio-evolutionary and socio-evolutionary time-scales to the variable one was not an abrupt and obvious transition; it was "a process like the coming of winter" [9].

As a result of the differing time-scales, i.e. changes in the social system occurring more quickly than our biological fitness could accommodate, reactions perfectly adapted to our old system remained with us even though the system for which they were adapted was gone. Our ingrained responses became less and less appropriate for our social system and humans were powerless to affect the development of more appropriate moral feelings. The coevolutionary Eden that guaranteed the moral fitting of our act to our society had vanished. The new order presented situations that our biology was unequipped to deal with, but not having any other recourse, the biological mechanisms produced responses based on what it had adapted to. Needless to say that mismatches and multiple matches occurred, triggering conflicting urges and socially unacceptable behavior.

Peter Singer tells a story about a woman Helen who must spend three months in Ecuador away from her serious boyfriend Bob. In Ecuador she has an affair with Juan. She tells Bob, truthfully, about her affair and that it is Bob whom she loves; Juan and Helen know that their connection will end when she returns to Bob. We fully expect Bob to react with jealousy; that is the reaction that biology has equipped us for in those situations. "Now that we have effective contraception, however, can't we put jealousy behind us?" Singer points out that feelings of jealousy are deep in our nature. There was a time when such feelings were exclusively beneficial, but those emotions are ill fit for the modern world. "[I]f our emotions were under more rational control" then we could eliminate our inappropriate intuitions and maintain moral feelings more in line with the demands of modern living. [14]

If the state of modern society is such as described, and or moral intuitions irreconcilably askew from what we need them to be, then how does the philosophical moral theory presented in the previous section apply to modern society? I have two ideas for approaches to connect the morality defined from the time when biological and social evolutionary times scales were identical to the time when they are not: accelerated bio-evolution and invented biology.

4.1 Accelerated Bio-Evolution

During the time when socio-evolution and bio-evolution ran at the same speed the moral truth of statements regarding individual actions, feelings, and social structure was guaranteed by the fact that those features coevolved. After the timescales diverged, the coevolutionary guarantee vanished. It is not that bio-evolution ceased, but merely that it could not keep up. Theoretically, if it could have kept up then our moral intuitions would still be the correct ones to have. So, to discover what feelings and responses are appropriate for modern society we should only need to uncover what feelings we would be left with if bio-evolution were accelerated to match socio-evolution.

The problem with this view is that the very process that we wish to harness to provide the appropriate feelings, the coevolutionary process, prevents us from uncovering what bio-evolution would equip us for if we sped it up. Both social and biological evolution adapt to each other, even as they progress at different rates. So by holding society constant and letting biology catch up we would be eliminating the necessary coadaptation ingredient from the formula. The biological base that we ended up with would not be the one that we would have had if they kept time together.

4.2 Invented Biology

The invented biology technique aims to create a biological base that preserves the appropriate behavior-situation relationships that we think are still relevant. It would not be valid to base such a biology on the moral intuitions that we now have about what is relevant for modern society because *ex hypothesi* those moral intuitions are sometimes incorrect. Determining the correct moral intuitions is exactly what we want the technique to tell us, so we cannot begin with any preconceived notions on the subject.

Alternatively we could examine the social structures of the protohuman society and the feelings and behaviors that they had within that society. Certainly whatever morality Mother Nature provided for the protohumans it was an appropriate one for their social structure. We could then construct a model of the salient relationships and relevant moral feelings and behaviors existing within that social structure and extend it to our modern social structure. Unfortunately, this would only tell us how animals like protohumans ought to act in a society like ours, not how we humans ought to act. We could follow the same process for wolves, for instance, but not think that we ought to act as wolves would act if the had to deal with irrationally jealous boyfriends.

4.3 Veil of Ignorance

Hence neither technique for extrapolating the known morally correct features of protohuman society to modern society is tenable. There may be other, better techniques, and future work will focus on attempting to uncover them. In the meantime we are left in the uncomfortable position that, according the theory presented above, statements of morality are either true or false but we have no reliable access to which is the case. We have emotive responses and elicited behaviors that may or may not be appropriate for our situation with no reliable guide as to whether they are or not. And the meanings of our moral terms, along with the concepts they identify or represent, may be distantly disconnected to our intuitions and muddled by conflicting responses to our modern environment.

We are not, however, fully without guidance towards correct moral attitudes. We can expect that since much of the human condition as not changed a great deal in the last 50,000

years or so, our moral intuitions will often be close to correct. We still need food, clothing, companionship, entertainment, etc. and even though technology has dramatically altered how we satisfy these needs, the drives we feel towards them perform the same function we needed them to as cave men. The fact that we cannot help but think they are correct is fully explained by the coevolutionary process that provided humans with the intuitions through biological means. That they might typically be correct would due to a happy coincidence that socio-evolution is not *that* much faster than bio-evolution. The theory presented here predicts, however, that things are only going to get worse.

5 Moral Models and Normative Ethics

Considering the hopeless inaccessibility of moral knowledge and helpless inability to alter our motivations imposed on modern humans according to the evolutionary moral theory presented above, what room is there for a normative moral theory? My answer is "none", but not because a system of action-guiding principles would be irrelevant or useless. Considering the reductionist, physicalist account of morality provided by my evolutionary framework the reason why a normative theory is impossible is because action-guiding principles could not constitute a *theory*. What is needed are models of society that identify what principles, if followed, tend to match our moral intuitions about what ought to be done.

To illustrate the distinction I will employ an example from physics and astronomy: Kepler's Laws. Kepler's three laws do not constitute a theory of planetary motion, they provide a system by which, given either an orbiting body's period of revolution or semi-major axis length, one can calculate the other. Combining these laws with data about the location and velocity of the planets he was able to construct a model of the solar system that correctly predicted one variable in terms of the other for all the observable planets. As new planets were discovered, they also fell well in line with Kepler's model's predictions (within a small degree or error). Through his model he was able to demonstrate that the orbits of planets must be ellipses instead of circles, contrary to the accepted belief of the time. Kepler's Laws, however, say nothing of *why* the orbits must be ellipses. Gravity, the supposed force generating that behavior, plays no explicit role in Kepler's model; nor need it. If we are interested in prediction and non-causal ex post facto explanation then as long as the underlying causal mechanisms remain mostly consistent, our model is sufficient for our purposes.

Kepler's model, however, is not a perfect match to observed phenomena. It abstracts away the effects the planets have on each other's orbits (among other things). Abstraction is a necessary feature of models; we cannot build exact duplicates of the phenomena that we wish to model because (among other reasons) doing so would remove the benefit of having the model. A necessary result of the model's simplification is that there will be special, unaccounted for, particular situations that do not fit the model. Often these anomalies result in better fitting models, other times they are simply accepted as being within the fault tolerance for our purposes of using the model.

A system of normative moral principles is analogous to Kepler's Laws for constructing a moral model of human society, with the evolutionary theory of motivations playing the role of gravity. What are some advantages for thinking of normative ethics as a mere model of the forces that humans are disposed to face with certain biologically determined attitudes? Well, models are not things that are correct or incorrect, but rather appropriate or inappropriate for one's purposes. The Ptolemy model of the solar system is not strictly speaking

wrong, but since its predictions are it would be inappropriate to use for satellite navigation. Models may, however, be better or worse than others with respect to being internally inconsistent, correlated to modeled phenomena, politically motivated, etc. And these kinds of considerations are precisely what many moral "theorists" debate about their normative "theories"- e.g. that they conform more tightly with our moral intuitions regarding various situations.

There are many different purposes to which one could commit a moral model. And for each purpose, a different moral model may be appropriate. The general list of condoned and censured actions remains quite consistent across normative moral models, as one would expect considering their aim to fit moral intuition. Each model has its own anomalous cases, and considering the complexity of the underlying causal framework (as described by the mismatched modern reactions to our more primitive moral feelings) it seems unlikely that one could construct a model fitting very closely.[13] By interpreting normative ethics as model building instead of theory construction, we must change our considerations of what to include and how to evaluate them.

6 Conclusions

I have presented as best as possible a complete moral framework for understanding morality in terms of our biological and social evolution and without reference to a faculty of free will. Putting aside the worry that the theory presented does not constitute a *moral* theory at all, what problems does the above theory fail to resolve?

6.1 Evolution as a Continuous Process

The theory developed here takes the appropriate response to be what had been selected for in the evolutionary history of the system in question. This imposes a consequentialist-style evaluation of the possible benefit of changes in the adapted behavior. Actual reactions are the right ones because if there were better reactions then those would have been selected for and the evolutionary process guarantees that these hypothetically better reactions would have been the actual reactions exhibited by agents within that system. But evolution never stops, and it depends on new reactions arising and being beneficial to generate the increasingly fit individuals and adaptive systems that it does.

The appearance of a new behavior that is inappropriate for a given system (in the sense that the system had evolved without any situation eliciting that behavior) may turn out to provide individuals with that new behavior a fitness advantage. The only way to determine whether such a mutations is beneficial, however, is to see how it plays out in the evolutionary process. But the evolutionary process never stops and there is always a small set of anomalous behaviors being tried and selected for or against. Because I wanted to differentiate the evolved systems at the two time-scale stages, I needed a formulation amenable to a static analysis. The counterfactual presentation of appropriate and inappropriate behaviors provides precisely that, but I lose any forward-looking capabilities.[14]

[13] Indeed, the forces regulating the motions of planetary bodies are quite simple compared to the tangled web of human-animal-environment interaction. Yet our best model of the solar systems continues to make egregious errors at the scales we are now concerned with.

[14] Note that consequentialist theories in general suffer from this problem. Utilitarianism, for example, takes that the aggregate satisfaction over states of affairs is the bearer of morality. But whether some act, institution, rule, etc. actually maximizes aggregate preference-satisfaction depends on how things turn out. There is no final state (hopefully) from which we can base our calculations.

An immediate result of adjusting the theory presented above to accommodate open-ended evolutionary processes is that at any given time, the truth or falsity of a moral claim is unknowableeven in the protohuman systems. This follows from that fact that determining whether a behavior is fitness enhancing depends on how events play out. The problem is analogous to cases of civil disobedience. Performing an action against the accepted norms of a society may be seen as reprehensible at the time, but if the behavior catches on then it is cast in the light progressive and innovative. For the protohumans I can identify a pseudo-final state, the last social system existing before socio-evolution accelerated beyond bio-evolution, but this is an artificial and unsatisfactory compromise. Accommodating open-ended evolution proved too difficult for the current work, but will be given serious consideration in the future.

6.2 Roles of Society and Psychology

6.2.1 The Nurture Argument

One strong argument against this theory would be that social factors are an important determiner for both moral feelings and other morally relevant emotions.[15] The possibility of social influence requires that there is a great deal of developmental variation and/or phenotypic plasticity in humans' emotional response mechanism. Remember that I am not talking about variations in moral rules (i.e. appropriate moral models), but variations in moral intuitions. According the theory here, it is not what we *think* is right or wrong, but rather what we *feel* is right or wrong (or react to in a way that we interpret as positive or negative) that is relevant.

6.2.2 Moral Development and Human Society

That humans' moral responses are learned through education, habituation, and socialization is one common explanation given for the variety of moral codes and institutions across the globe. One might compare internalizing a moral code (and hence developing a response mechanism coherent with that code) to learning a language. Just as few individuals can tell you the rules of grammar for their language yet can speak generally in accordance to them, so too can people act morally without knowledge of the structure of their moral system. Our genetic make-up may include the template for the capability of learning a linguistic or moral system, but neither system comes hard-wired. Humans may not have evolved with specific hard-wired moral feelings, but rather the ability to adapt our moral feeling to our environment.

6.2.3 Role of Free Will

In the modern era, we experience conflicts of moral feeling that could not arise in protohumans. Because of the mismatch of the biological mechanism for generating emotions and the social situations that elicit them, human interactions will sometimes trigger multiple moral feelings. If the feelings are of different valences then we have a conflict of moral intuitions. The biological mechanisms in place are not capable of dealing with such conflicts because they did not arise at the time when biological evolution could keep up with social evolution. Such conflicts cause much stress and cognitive dissonance in modern society. Regulating our emotional reactions and choosing which one to act upon is precisely the role

[15] Recall from above that I accept the possibility that not all emotional responses are relevant to the moral structure, though I see no argument at present to exclude any emotion from being possibly morally relevant.

for which we commonly deploy the faculty of conscious free will in psychology and moral theories. If sufficient evidence can be shown that our mental processes do, in fact, perform such a function then intentional action would save modern humans from the dismal position my account puts them in. Though such a capability would be a welcome addition to our response regulatory mechanism, I am afraid that believing in conscious free will is merely wishful thinking. Wanting it, or even needing it, doesn't make it so.

6.2.4 Future Work

These issues (and others that will certainly arise) shall be considered in future work. Variations on the underlying framework can provide for alternative ideas regarding the roots of our moral experiences and how we react to them. Empirical studies on the brain's mechanisms will no doubt provide further insight into the relationships that we have to our environment and their results can bolster an evolutionary account of our moral systems. It would also be valuable to more thoroughly compare existing moral theories to the theory presented here and determine to what degree they match and/or contrast with its conclusions. As new avenues are considered we can better determine whether an ethical theory grounded on evolutionary biology is possible and gain a better understanding of our moral foundations through the development of these ideas.

References

[1] Blackmore, Susan. *The Meme Machine* (Oxford: Oxford University Press, 2000).

[2] Brink, David. *Moral Realism and the Foundations of Ethics* (Cambridge University Press: 1989).

[3] Boyd, Richard, "How to Be a Moral Realist" excerpted in *Moral Discourse & Practice* (New York: Oxford University Press, 1997), pp. 105-136. First published in 1988.

[4] Crick, Francis and Christof Koch. "Consciousness and Neuroscience" in *Cerebral Cortex* 8 (1998), pp. 97-107.

[5] Davidson, Donald. "Mental Events" in *Essays on Actions and Events* (Oxford: Clarendon Press, 1980), pp. 207-224. First published in 1970.

[6] Dawkins, Richard. *The Selfish Gene* (Oxford: Oxford University Press, 1976).

[7] Dennett, Daniel. *Freedom Evolves* (New York: Penguin Group, 2003).

[8] Hairston Jr, Ellner, Geber, Yoshida and Fox. "Rapid evolution and the convergence of ecological and evolutionary time" in *Ecology Letters* (2005) 8: 1114-1127.

[9] Hare, Richard. *The Language of Morals* (London: Oxford University Press, 1961).

[10] Kant, Emmanuel *Groundwork for the Metaphysics of Morals* Translated by Lewis Beck. (Indianapolis: Bobbs-Merril Company, Inc., 1969).

[11] Kim, Jaegwon. "Concepts of Supervenience" in *Philosophy and Phenomenological Research* Vol. 45, No. 2. (Dec., 1984), pp. 153-176.

[12] Machiavelli, Niccolo. *The Discourses*Translated by Daniel Donno. (New York: Bantam Books, 1966).

[13] Nagel, Thomas "The Possibility of Altruism" excerpted in *Moral Discourse & Practice* (New York: Oxford University Press, 1997), pp. 323-340. First published in 1970.

[14] Singer, Peter. "On Jealousy" in *The Daily Princetonian* (October 6, 2005). Available at www.utilitarian.net/singer/interviews-debates/20051006.htm

[15] Smith, Michael. *The Moral Problem* (Malden: Blackwell Publishing Ltd, 1994), pp.124-181.

[16] Wegner, Daniel. *The Illusion of Conscious Will* (Cambridge: MIT Press, 2002).