The Essence of Morality: A Teleological-Formalist Conception of Moral Progress

Daniel L. Zhang
University of Pennsylvania

5-1-2015

Follow this and additional works at: http://repository.upenn.edu/ppe_honors

Part of the Behavioral Economics Commons, Behavior and Behavior Mechanisms Commons, Ethics and Political Philosophy Commons, Social Psychology Commons, and the Theory and Philosophy Commons

The Essence of Morality: A Teleological-Formalist Conception of Moral Progress

Abstract
This thesis considers the problem of how people should deal with moral change. I examine various conceptions of morality and induce the essence of morality, or the set of formal features that is fundamental to all moralities. In particular, I claim that the essence of morality consists of four features: (1) it performs some function (namely, solving the problem of cooperation); (2) it is prescriptive; (3) it is regarded as universalized; and (4) it is regarded as possessing overriding authority over other normative systems. From this, I derive a metamorality or standard of moral progress with which to evaluate moral change, namely the teleological-formalist conception of morality, which holds that morality progresses if and only if it increasingly exhibits its formal features. This is important because it allows philosophers and policymakers more generally to evaluate moral change and steer moral progress.

Subject Categories
Behavioral Economics | Behavior and Behavior Mechanisms | Ethics and Political Philosophy | Social Psychology | Theory and Philosophy

This article is available at ScholarlyCommons: http://repository.upenn.edu/ppe_honors/6
THE ESSENCE OF MORALITY:
A TELEOLOGICAL-FORMALIST CONCEPTION OF MORAL PROGRESS

Daniel L. Zhang

A THESIS
in
Philosophy, Politics, and Economics

Presented to the Faculties of the University of Pennsylvania
in
Partial Fulfillment of the Requirements for the
Degree of Bachelor of Arts with Departmental Honors

2015

Supervisor of Thesis
Co-Supervisor of Thesis

Jan Willem Lindemans
Sumantra Sen
Postdoctoral Fellow
Associate Director of
Undergraduate Studies
THE ESSENCE OF MORALITY

COPYRIGHT

2015

Daniel Ling Zhang

This work is licensed under the
Creative Commons Attribution-
NonCommercial-ShareAlike 3.0
License

To view a copy of this license, visit

http://creativecommons.org/licenses/by-ny-sa/2.0/
To my parents

For giving everything and more
ACKNOWLEDGMENTS

I would like to express my deep gratitude to Dr. Jan Willem Lindemans for his continual support and guidance throughout the semester. I am grateful for his recommendations, contributions, and expertise, without which this project would not have been possible. I would also like to thank Dr. Sumantra Sen, who sparked my interest in writing a thesis.
ABSTRACT

THE ESSENCE OF MORALITY

Author: Daniel L. Zhang
Supervisor: Jan Willem Lindemans

This thesis considers the problem of how people should deal with moral change. I examine various conceptions of morality and induce the essence of morality, or the set of formal features that is fundamental to all moralities. In particular, I claim that the essence of morality consists of four features: (1) it performs some function (namely, solving the problem of cooperation); (2) it is prescriptive; (3) it is regarded as universalized; and (4) it is regarded as possessing overriding authority over other normative systems. From this, I derive a metamorality or standard of moral progress with which to evaluate moral change, namely the teleological-formalist conception of morality, which holds that morality progresses if and only if it increasingly exhibits its formal features. This is important because it allows philosophers and policymakers more generally to evaluate moral change and steer moral progress.
# TABLE OF CONTENTS

ACKNOWLEDGMENTS ........................................................................................................... iii

ABSTRACT ................................................................................................................................. iv

INTRODUCTION ........................................................................................................................... 6

PART ONE: THE ESSENCE OF MORALITY .............................................................................. 10  
*The essence of morality* ........................................................................................................ 12  
(1) Morality pursues some function .................................................................................... 13  
(2) Morality is prescriptive .................................................................................................. 20  
(3) Morality is regarded as universalized ............................................................................ 20  
(4) Morality is regarded as possessing overriding authority over other normative systems ............................................................................................................................ 21

PART TWO: THE TELEOLOGICAL-FORMALIST CONCEPTION ............................................. 23  
*Moral change and progress* .............................................................................................. 24  
*Deriving metamorality from the function of morality* ......................................................... 26  
*Necessary conditions for solving the problem of cooperation* ........................................... 29  
Prescriptiveness .................................................................................................................... 29  
Perception of universality ..................................................................................................... 30  
Perception of overriding authority over other normative systems ...................................... 30

PART THREE: IMPLICATIONS ............................................................................................... 32  
*Metamoral judgments: evaluating moral change* ............................................................... 32  
*Steering moral progress using the teleological-formalist conception* ............................... 34

CONCLUSION ........................................................................................................................... 36

BIBLIOGRAPHY ......................................................................................................................... 38
INTRODUCTION

Recently, scholars have investigated and debated the nature of moral change. For example, the emerging study of the evolution of cooperation has captivated a diverse range of scholars in biology, psychology, economics, and political science. However, while much effort has been dedicated to the empirical study of moral change, the normative questions regarding moral change remain largely unanswered.

In this thesis, I considered the following question: if it indeed exists, how should we deal with moral change? Phrased differently: how should we evaluate or normatively relate to moral change? Some scholars, especially sociologists and historians, have identified and evaluated moral change in the past, providing ex post assessments of moral change. My goal, however, is to investigate the possibility of conducting ex ante assessments of moral change: that is, evaluating moral change before and while it occurs. In doing so, I attempt a task that philosophers and scholars more generally have been largely unwilling to attempt.

In this paper, I set forth a standard of moral progress with which to evaluate moral change. This standard can be understood as a “morality of moralities” or metamorality, in that it is a guide for making moral judgments about moralities as they change with time. This standard can thus be used to make ex ante assessments about moral change.

To arrive at this standard, I attempt to abstract an essence of morality. The essence of morality is the set of formal characteristics shared by all moralities throughout history. It is analogous to the notion of artistic style, or the set of formal characteristics shared by all works of a specific painter. Just as one induces an artistic style from the common characteristics of a set of
paintings, so too do I attempt to induce the essence of morality by examining the common features of various conceptions of morality. From this induction, I derive a metamorality or standard of moral progress with which to evaluate moral change. Returning to the painting analogy, knowledge of a painter’s style or essence allows experts to identify forgeries; this ensures that new discoveries are correctly evaluated. Similarly, knowledge of the essence of morality allows philosophers to judge moral discoveries. Moreover, under this standard, philosophers and policymakers more generally can evaluate and steer moral change in a positive direction. This becomes particularly relevant as our society confronts complex and difficult issues, such as abortion and physician-assisted suicide, which may be a sign of impending moral change.

In my analysis, I find that the essence of morality consists of four formal features:

(1) It performs some function.
(2) It is prescriptive.
(3) It is regarded as universalized.
(4) It is regarded as possessing overriding authority over other normative systems.

Phrased differently, the essential morality *pursues a particular function* by *prescribing* rules or principles regarded as *universal* and *possessing overriding authority over other normative systems*. Moreover, I argue that the particular function pursued by morality is solving the problem of cooperation.
Using this essence, I derive a metamorality with which to evaluate moral change, namely the *teleological-formalist* conception of moral progress. The conception is *teleological* in holding that morality ought to aim towards some goal; and *formalist* in holding that moral progress is the refinement of the essence or formal elements of morality. Therefore, the conception is *teleological-formalist* in holding that morality ought to strive towards perfecting the formal elements of morality. More precisely, according to my teleological-formalist conception, morality progresses if and only if:

(1) It increasingly serves its function (namely, solving the problem of cooperation).
(2) It is increasingly prescriptive.
(3) It is increasingly regarded as universalized.
(4) It is increasingly regarded as possessing overriding authority over other normative systems.

This thesis is structured as follows. In Part One, I induce the essence of morality from various conceptions of morality. First, I present Frankena’s (1966) formal elements of morality, which are representative of the scarce literature on metamorality, and argue that this account is insufficient. In response, I propose and defend four features (listed above) that I believe compose the essence of morality. Furthermore, I argue that the function of morality is to solve the problem of cooperation.

In Part Two, I use this essence to derive a conception of moral progress, or “metamorality,” with which to evaluate moral change. In particular, I examine various
conceptions of moral progress and propose the teleological-formalist conception of moral progress, which holds that morality progresses if and only if it increasing exhibits the formal features of morality.

In Part Three, I conclude by considering the implications and practical uses of this proposed conception. In particular, I briefly describe how philosophers and policymakers more generally can use the teleological-formalist conception to evaluate moral change and steer moral progress.
PART ONE: THE ESSENCE OF MORALITY

To evaluate moral change, I will abstract the essence of morality from different moralities and use this essence as an evaluative criterion. In this part, I consider what features are included in the essence of morality. I begin my analysis with Frankena’s (1966) conception of formal morality, which I take to be roughly representative of the scarce literature on the essence of morality. I find that Frankena’s conception is inadequate and offer my own account of the essence of morality. In the remainder of the first part, I justify and defend these features from objections, particularly from social intuitionist psychologists and moral relativist philosophers.

Before proceeding, it is important to consider whether or not morality has an essence. An essence consists of the fundamental features of some thing that allow it to be distinguished from other things. More precisely, an essence of \( X \) is the smallest set of features that can be used to define and identify \( X \). For example, the essence of a chemical element consists of one feature, namely that it is composed of a single kind of atom. This allows us to distinguish a chemical element from a chemical compound, which consists of two or more different chemical elements that combine through chemical reaction, or a chemical mixture, which consists of two or chemical elements that do not combine through chemical reaction.

Morality also has an essence because it can be readily distinguished between other things. For example, we are able to distinguish between moral obligations and other types of obligations, such as legal or cultural obligations. Moreover, moral reasoning feels different from other forms of reasoning, such as logical reasoning; logical reasoning, for example, rarely arouses guilt or
shame, whereas moral reasoning can and often does arouse these emotions. Because we are able to distinguish between morality and other similar things, there must be some set of fundamental features that defines and constitutes the essence of morality.

However, few philosophers have attempted to induce the essence or formal elements of morality. Frankena’s (1966) proposed set of formal features is roughly representative of the scant literature on the topic (e.g., Hare, 1952; Hare, 1963; Gert, 2005). Frankena proposes that formal morality consist of three features; in particular, a normative system is morality if and only if it is:

1. prescriptive;
2. universalized; and
3. possesses overriding authority over other normative systems (such as legal systems).

(1) is an uncontroversial feature: prescriptiveness is a necessary part of any normative system and therefore must be a feature of morality. (2) and (3) are more contentious. In particular, social intuitionist psychologists and moral relativists might argue that morality is a post hoc rationalization of cultural norms or self-interested goals and is therefore neither universalized nor supremely authoritative (e.g., Haidt, 2001; Knobe & Leiter, 2007).

Overall, Frankena’s conception is unsatisfying because it does not explain why morality exists. Moreover, on the surface, it is unclear why the features are essential to morality: for example, why must morality be prescriptive? Frankena’s answer seems to be: if morality is not prescriptive, then it would not be morality. To avoid this circularity, we need a formal conception of morality that explains why morality and its formal features exist.
Moreover, Frankena’s conception is insufficient for the purposes of deriving a metamorality. It is unclear how these features can be used to anticipate and evaluate moral change. One possibility might be that we can say that moral change is good if morality increasingly exhibits its formal features and bad if it does not. However, using Frankena’s conception, this would lead to absurd conclusions: for example, moral change towards a system that universally prescribes selfish behavior with absolute authority over other domains would be seen as good. Therefore, Frankena’s conception is inadequate for purposes of evaluating moral change.

**The essence of morality**

I propose a revision to Frankena’s conception and suggest that morality consist of four formal features:

(1) It performs some function.

(2) It is prescriptive.

(3) It is regarded as universal.

(4) It is regarding as having overriding authority over other normative systems.

(1) is a new feature that explains why morality exists, which was missing in Frankena’s account. Namely, morality exists in order to perform a particular function.

(2), (3), and (4) are revisions of Frankena’s existing features. (2) preserves in its entirety Frankena’s feature of prescriptiveness. (3) and (4) make minor but consequential adjustments:
from “universalized” to “regarded as universalized,” and from “possessing overriding authority” to “regarded as possessing overriding authority” (respectively). These adjustments solve the problems described earlier based on social intuitionist and moral relativist objections; in particular, (3) and (4) concede that while morality may not actually be universalized or overriding authority, it is nonetheless true that people regard or behave as if morality were universalized and overriding authority.

Moreover, as I will later demonstrate, (1) implies (2), (3), and (4). In other words, morality cannot pursue its function without being prescriptive, regarded as universalized, or regarded as possessing overriding authority over other normative systems. In my view, these four features together comprise the essence of morality.

(1) Morality pursues some function

Evolutionary accounts of morality

Evolutionary accounts of morality posit that morality serves the function of maximizing fitness. Waddington (1957) defines fitness as “the capacity to contribute offspring to the next generation” (p. 109). Similarly, Mayr (1991) defines fitness as “[t]he relative ability of an organism to survive and transmit its genes to the gene pool of the next generation” (p. 189). Dennett (1995) characterizes this definition in more detail by proposing that $x$ is fitter than $y$ if and only if $x$’s traits allow it to survive environmental pressures and reproduce more fully than do $y$’s traits. Fitness can also be defined in an absolute or relative sense. Absolute fitness considers the total fitness of a particular population, while relative fitness is a comparative
measure in which absolute fitness is normalized across genotypes within a population (Orr, 2009).

Defining morality in relation to evolutionary fitness seems counterintuitive. Moral behaviors typically do not lead to better chances of survival and reproduction. Consider the case of cooperation. Cooperation does not appear to be advantageous in terms of fitness (see generally West, El Mouden, & Hardner, 2011). All other things being equal, cooperation seems to reduce the relative fitness of a cooperating individual while increasing the relative fitness of the recipients of the cooperative act. Therefore, cooperative behavior would be weeded out by natural selection. Non-cooperating individuals would free ride off the cooperative behaviors of others, increasing the fitness of the free riders at the expense of the cooperators. As a result, genes for free riding would have greater fitness and be retained in future populations. This seems to imply that cooperation, and morality in general, cannot be explained in terms of fitness.

Scholars have proposed at least two solutions to the problem presented above. First, Hamilton’s (1964) kin selection theory posits that altruistic genes can be selected by evolution due to kin-cooperative relationships, in which an individual cooperates or assists a closely-genetically-related individual. Even if altruism is costly for the altruistic individual, the benefits of altruism will be imparted onto an individual with similar genes to the altruistic individual. Altruism thus results in increased fitness of the recipient, who shares genes with the altruistic individual; therefore, this increases the fitness of altruistic genes shared between the two kin-related individuals. This “indirect” passing of genes (i.e. from an altruistic individual via fitness gains received by a closely-genetically-related individual) gives way to a new measure of fitness, namely inclusive fitness, or personal fitness plus fitness afforded via altruism to kin-related
individuals. Kin selection thus explains morality in terms of fitness: while moral behaviors, such as altruism, may decrease the personal fitness of the individual, they tend to increase the inclusive fitness of the individual. Natural selection, then, tends to maximize of the inclusive fitness of individuals in a population (see Grafen, 2006). Indeed, nature is abundant with examples of altruistic behavior that seems to have evolved through kin selection (e.g., Queller & Strassmann, 1998; Russell & Hatchwell, 2001; Morin, Moore, Chakraborty, Jin, Goodall, & Woodruff, 1994; see generally Clutton-Brock, 2002).

Despite this, kin selection does not provide a comprehensive explanation of morality. The theory implies that altruism becomes more likely as individuals are more closely related, since higher genetic similarity between individuals means a higher probability that the individuals share altruistic genes to be passed onto future generations (Hamilton, 1972). However, this does not explain altruism between non-genetically-related individuals, which is observed in nature and often prescribed by morality (e.g., Ligon & Ligon, 1990; Reyer, 1990; Zahavi, 1990; Brosnan & de Waal, 2002; Schino & Aureli, 2009).

Reciprocal altruism provides an explanation of altruism between both genetically-related and non-genetically-related individuals. Reciprocal altruism theory posits that under repeated interaction between individuals, altruistic genes will be favored if the cost of altruism to an altruistic individual is exceeded by future reciprocal benefits “returned” by the recipient of altruism (Trivers, 1971). Axelrod and Hamilton (1981) provide a useful illustration of reciprocal altruism using game theory. While cooperation is unlikely in a one-shot or single iteration of the

---

1 Formally, kin selection theory holds that altruistic genes will be favored if Hamilton’s rule is satisfied (Hamilton, 1964). Hamilton’s rule is \( b > c/r \), where \( b \) is the benefit to the recipient of altruism, \( c \) is the cost to the altruistic individual, and \( r \) is the coefficient of relationship between the actor and recipient, which is a measure of genetic relatedness. As \( r \) increases, the probability that Hamilton’s rule will be satisfied tends to increase.
prisoner’s dilemma, it becomes more likely as the game is played repeatedly. Suppose that players adopt a tit-for-tat (TFT) strategy, such that in any present round, each player mirrors what the other player chose in the previous round. This strategy is conducive to reward (for cooperation), punishment (for defection), and forgiveness (for players who change from defectors to cooperators). TFT therefore uses the expectation of reciprocity to encourage cooperation and discourage defection. Individuals in nature may use a similar strategy in choosing whether or not to engage in altruistic behavior. Relationships governed by reciprocity, then, may allow altruistic genes to be favored by natural selection.

While reciprocal altruism theory explains moral behaviors between non-genetically-related individuals, it is still insufficient in explaining morality more generally. For example, the theory explains altruism conditioned by reciprocity but does not explain unconditional altruism, which seems to be a commonly prescribed by moral systems. As a result, reciprocal altruism seems to “take the altruism out of altruism” (Trivers, 1971, p. 35).

**Non-evolutionary accounts**

Alternatively, some scholars have offered non-evolutionary accounts of morality. Posner (1999), for example, contends that morality serves a regulatory function, or that morality functions as: “the set of duties to others…that are supposed to check our merely self-interested, emotional, or sentimental reactions to serious questions of human conduct” (p. 4). Similarly, Mackie (1977) suggests that morality solves the problem of “limited resources and limited sympathy,” or problems arising from scarcity of resources and flawed human dispositions that fail to promote fair or efficient allocations of resources. Copp (2009) explains morality as a
solution to “problems of normative governance,” which impose limits on our ability to achieve what we value (p. 27). In particular, morality solves the problems of sociality and politeness by facilitating cooperation and coordination (see also Alexander, 1987; Krebs, 2000). The explanation of morality as a coordinating mechanism is what Lillehammer (2003) calls the “standard” view of the function of morality (p. 569).

Gauthier (1986) provides a particular instantiation of the standard view that is worth investigating, known as the “morality-as-rationality” view. Gauthier begins from the assumption that morality is revealed through reason. Drawing upon Hobbesian political philosophy, Gauthier supposes a state of nature in which people face problems of whether or not to cooperate with others. Gauthier argues that, in the absence of credible enforcement mechanisms, human instincts, as well as uncertainty about others’ intentions, encourage people to shirk cooperation. However, rationality reveals that cooperation would have provided a larger overall endowment and is therefore the optimal strategy.

Up until this point, Gauthier’s conceptualization mirrors that of Hobbes; however, whereas Hobbes proposes a strong sovereign to solve the problem of enforcement, Gauthier argues that humans, as rational beings, create morality. Morality begins as a set of basic obligations, which includes norms of promise-keeping and trust, to facilitate cooperation in situations of uncertainty. Therefore, Gauthier argues, morality is formed as a “rational constraint on the pursuit of individual interest” (1986, pp. 1-2). Sidgwick (1981) echoes this view, arguing that morality is an element of the “dualism of practical reason”: on the one hand, practical reason is self-oriented, in that it is characterized by prudential self-interest; on the other hand, practical reason is socially- or other-oriented, in that it prescribes moral obligations to other individuals (p.
xii). Morality regulates the conflict between the self-oriented and other-oriented aspects of rationality.

**A hybrid account**

Gauthier’s view seems to require a strong assumption, namely that primitive humans, when confronted with the problem of cooperation, were sufficiently rational to realize and implement morality as a solution. However, Trivers’ theory of reciprocal altruism offers a solution to this problem. Let us concede the previous objection; that is, suppose that primitive beings were not sufficiently rational to realize that morality served a useful function. Even so, reciprocal altruism theory requires only that those beings that did create and engage in moral behaviors were more likely to survive and reproduce than those that did not. For example, those who formed moral relationships to enforce cooperation were more likely to form large and tight-knit groups with greater inclusive fitness than those who shirked cooperation (see Clutton-Brock, 2009). The theory does not require beings to be rational nor does it require beings to realize that morality is rational. In fact, it is likely that those who instinctively or even irrationally adopted morality were most likely to survive, because unconditional moral obligations tend to be strongest and thus most effective in enforcing cooperation, promise-keeping, and relationships of trust (see Lotem, Fishman, & Stone, 2003; Sampson, 2003).

This also addresses with the objection that reciprocal altruism is not really altruism, or that it takes the “altruism out of altruism” (Trivers, 1971, p. 35). The theory only requires that altruism occurs under conditions of reciprocity and does not require that actors are cognizant of the reciprocal nature of their altruistic relationships. Therefore, reciprocal altruism theory
explains how people may have evolved unconscious moral behaviors, which are ultimately driven by reciprocity. Reciprocal altruism theory, and evolutionary theory more generally, therefore offers a plausible scientific explanation for the standard view. This particular explanation, which uses evolutionary theory to illustrate philosophical claims, can be called the hybrid view.

**The function of morality**

I contend that the function of morality is to solve problems of cooperation. This claim is supported by both evolutionary accounts, as in the theory of reciprocal altruism, as well as non-evolutionary accounts of morality, as in the standard view. If morality did not serve this function, it would be difficult to imagine how morality could have developed at all. While there are other solutions to the problem of cooperation, such as laws, government, and culture, these normative systems generally flow from morality. For example, we tend to think of laws as a set of rules that reflect some but not all elements of morality more generally. Similarly, cultural practices are intimately linked with and sometimes inseparable from moral obligations (see generally Haidt, 2008). Moreover, moral authority tends to override legal or cultural authority (see Fuller, 1965). As a result, morality seems to be the most basic and prototypical solution to the problem of cooperation.

Moreover, moralities are no more than various solutions to the problem of cooperation. Consequentialist systems, for example, solve the problem of cooperation by emphasizing the good consequences brought about by cooperation. Deontological systems demand respect for rules such that cooperation becomes a duty. Finally, virtue ethics emphasizes the importance of good character such that the tendency to cooperate is seen as a sign of virtue. In other words,
each system aims towards the same end but differs in its proposed means. Therefore, I contend that the function of morality more generally is to solve the problem of cooperation.

(2) Morality is prescriptive

Prescriptiveness is an uncontroversial part of Frankena’s model. In particular, prescriptiveness is a necessary feature of all normative systems; thus, morality, as a normative system, must be prescriptive. Therefore, Frankena’s feature of prescriptiveness has been preserved in this revised conception.

(3) Morality is regarded as universalized

Frankena’s claim that morality is universalized can be criticized on the grounds that moral reasoning, judgments, and behaviors are not uniform. Moreover, moral relativists reject the metaphysical notion of a universal morality and emphasize the intense moral disagreements between groups and cultures (see generally Krausz, 2011). Similarly, psychologists, particularly those who subscribe to social intuitionism, contend that universal morality does not exist because moral reasoning is the product of post-hoc rationalization; that is, people have emotional or instinctual responses to certain stimuli, and then use moral reasoning to justify those responses (e.g., Haidt, 2001; Haidt, 2007; Tetlock, 2002). Thus, morality is highly situational, which implies that there cannot be a universal morality.

It is more plausible to say that morality is regarded as universalized. The previous objections make the descriptive claim that morality is not universal. Even so, it is true that people interpret morality as universal. That is, whether or not morality is universal in a metaphysical sense, people tend to reason and behave as if morality were universal (cf. Davie, 1988). This
would explain the intensity of moral disagreements often cited by moral relativists as evidence against universal morality: such disagreements are intense because each party believes that its own morality is universal and thus ought not be challenged. Therefore, an accurate conception of morality ought to include the feature that morality is regarded as universalized.

(4) Morality is regarded as possessing overriding authority over other normative systems

Frankena’s claim that morality possesses overriding authority over other normative systems can be criticized on similar grounds as those described in the previous subsection. For example, social intuitionists may argue that culture determines affective responses to stimuli, which in turn determine our moral reasoning through post hoc rationalization; therefore, culture actually has overriding authority over morality (see Haidt & Joseph, 2007; Shweder, 1982).

Even so, it remains true that people behave as if morality possesses overriding authority over other normative systems (cf. Kar, 2012). For example, there are no laws against lying to friends and family, but we generally agree that this form of lying is bad, regardless of our culture. That is, moral obligations transcend legal or cultural obligations. Therefore, we can say that morality is regarded as possessing overriding authority over other normative systems.

In this part, I have refined Frankena’s conception of formal morality and arrived at the following conception, which I describe as the essence of morality; namely, that morality consists of the following essential or formal features:
(1) It pursues a particular function.
(2) It is prescriptive.
(3) It is regarded as universalized.
(4) It is regarded as possessing overriding authority over other normative systems.

In the next part, I will use this essence to derive a standard of moral progress or metamorality with which to evaluate moral change.
In this part, I will consider the notion of moral progress. The purpose of inducing the essence of morality in Part One is to use it as a standard or “metamorality” with which to evaluate moral change. Thus, what we are really after is a standard of moral progress that tells us whether moral change is good or bad. For the purposes of this paper, I will use the terms “metamorality” and “standard of moral progress” interchangeably.

In my analysis, I will conclude with what I call the teleological-formalist conception of moral progress. It is teleological in supposing that moral progress aims towards some goal, namely solution of the problem of cooperation; and formalist in supposing that achievement of this goal relies on perfecting the essence or formal features of morality, namely the features I induced in Part One. More precisely, the teleological-formalist conception of moral progress states that morality progresses if and only if it increasingly exhibits the following features:

1. It serves some function (namely, solving the problem of cooperation).
2. It is prescriptive.
3. It is regarded as universalized.
4. It is regarded as possessing overriding authority over other normative systems.
I begin by considering several conceptions of moral progress proposed by various scholars. I conclude that these conceptions are logically reducible to the teleological-formalist conception stated above.

**Moral change and progress**

Morality is not static; rather, it has been subject to significant movements throughout history. Consider the fact that just two centuries ago, the United States condoned and even supported slavery; now, we find it unquestionably repugnant. According to evolutionary theory, this rate of change is unsurprising: typically, the rate of evolution of a certain trait is determined by the rate of reproduction and the amount of variability in genotypes of that trait (see generally Smith, 1976). In particular, a high rate of reproduction and/or low genetic variability tends to lead to a faster rate of evolution (Hartl & Clark, 1997). While humans reproduce more slowly than do other living things, such as bacteria, the variability of human genotypes in 19th century America (at the time of slavery), a largely homogenous society, was low (see Short, 1976; Durham, 1991; *e.g.*, Genovese, 1965). Therefore, we would expect to see rapid and pronounced change in moral behavior and reasoning. Indeed, studies in the field of value research provide persuasive evidence of rapid moral change. Inglehart’s work, for example, has revealed significant change from materialist values towards postmaterialist values in only the past forty years (1995; Inglehart & Abramson, 1994; Inglehart & Baker, 2000). Therefore, moral change does in fact occur.

This raises the question: how should we evaluate moral change? Sometimes, moral change is good: for example, we generally agree that our changing attitudes toward slavery and women’s rights are positive changes. On the other hand, there are instances in which moral
change is bad. For example, some sociologists and historians have criticized moral values that emerged in the Industrial Age, identifying them as signs of moral decline (e.g., Spengler, 1922/1932; Weber, 2013/1922). In both these examples, moral change is evaluated retroactively: given the benefit of hindsight, scholars consider whether moral change was positive or negative. However, an optimal approach would be to evaluate moral change as it occurs, without the benefit of hindsight. In doing so, we might be able to steer moral change away from bad outcomes and towards good outcomes. In order to do so, we need a conception of moral progress, or some way of knowing whether change is morally progressive or regressive.

Scholars have offered many conceptions of moral progress, which can be divided into at least two categories. The first category considers the content of moral progress: that is, what goals of principles does moral progress consist in. Jamieson (2002) defines moral progress as the extent to which “a subsequent state of affairs is [morally] better than a preceding one, or when right acts become increasingly prevalent” (p. 20). Godlovitch (1998) suggests that moral progress occurs when moral obligations cover an increasing number of domains and govern each domain more strictly. Macklin (1977) argues that moral progress involves increased application of the “principle of humanity” (sensitivity to pain and suffering) and “principle of humanness” (sensitivity to human dignity and rights) (p. 371).

The second category of definitions investigates the process of moral progress. That is, this category describes the particular path of moral change that consists in a progressive moral society. Wilson (2010) argues that moral progress is similar to scientific progress, in that both involve movement along “unidirectional narratives” (p. 97). Movement of progress follows a unidirectional narrative if movement from one belief to another belief is seen as progressive,
while the reverse is seen as regressive. This is contrasted with bidirectional narratives, in which
the movement from one belief to another belief is seen as neither progressive nor regressive;
rather, the consideration of an increasing number of beliefs in itself seen as progressive, whereas
the reverse is regressive. Bidirectional narratives, therefore, progress towards pluralism of moral
belief systems. Moral progress, Wilson argues, follows a unidirectional narrative in which there
is a clear sense of progress or regress simply due to the direction of belief change.

**Deriving metamorality from the function of morality**

Each of the previous conceptions seems to be reducible to a single form: that morality
progresses if it increasingly serves its function. For example, consider Macklin’s conception,
which contends that morality progresses as it increasingly applies the principles of humanity
(sensitivity to pain and suffering) and humaneness (sensitivity to human rights and dignity). The
two principles refer to consequentialist and deontological principles, respectively. However, as
demonstrated in the Part One, both consequentialism and deontology, and indeed all moral
philosophies, are attempts to solve the problem of cooperation. That is, they share a common
function. Therefore, Macklin’s conception can be stripped down to a conception holding that
moral progress occurs as morality increasingly serves its function, or increasingly solves the
problem of cooperation.

The function of morality has normative importance. Consider that morality must serve
this function in order to persist. For example, according to evolutionary accounts of morality
discussed in Part One, morality serves the purpose of increasing genetic fitness. If morality were
to fail in performing this function, genotypes that code for moral behaviors and cognitions would
gradually fade in future populations. As a result, we would no longer observe moral behaviors in the long run. Similarly, in the non-evolutionary accounts of morality, morality performs a function that is intimately linked with survival. In particular, morality is a necessary means of solving problems of cooperation. Supposing that cooperation is essential for human survival, this means that morality, by transitivity, is itself a necessary condition for human survival (see Talmage, 1993-1994). For example, consider a world in which morality fails to solve the problem of cooperation. As a result, human suffering is unchecked and human rights are nonexistent. It would be nearly impossible for large groups and societies to form: humans, without the guarantee of cooperation provided by moral obligations, would shirk cooperation and remain in small, weak groups. Without biological or physiological advantages afforded to other animals, humans would thus be driven from their ecosystems and exposed to intense survival pressures (see generally Neuberg & Cottrell, 2008). As human populations diminished, so too would moral obligations and behaviors. This can be conceptualized as a feedback loop through which morality and survival are mutually maintained: as morality diminishes, survival diminishes; as survival diminishes, so too does morality. Therefore, morality tends to disappear as morality increasingly fails to serve its function.

Therefore, I claim that morality ought to serve its function. That is, the function of morality has normative importance; thus, moral progress involves the successful performance of this function. If morality failed to serve its function, then morality would disappear, whether through the disappearance of its genotypes from human populations or the death of human populations that practice and shape morality. That is, the function of morality is a critical and
existential component of morality. Therefore, we ought to care about whether or not morality serves its function.

Thus, I conclude that a proper conception of moral progress ought to be defined in relation to the function of morality, namely solving the problem of cooperation. This appeals to our intuitions: if values and principles change such that we are more likely to cooperate, we see this as progressive. However, one might object that this conception leads to absurd conclusions. For example, slavery actually led to cooperative outcomes: slaves and slaveholders seemed to be in a state of (forced) cooperation. As we moved away from slavery, cooperation became less prevalent: for example, the abolition of slavery in the United States was followed by a period of racial violence and institutionalized prejudice. If we accept this conception of moral progress, this seems to suggest that our changed attitudes towards slavery were morally regressive, which is surely incorrect. Thus, the objection concludes, moral progress cannot be defined in terms of increasing cooperation.

However, in this conception of moral progress, “cooperation” refers to the problem of cooperation framed in the earlier discussion on reciprocal altruism. That is, the problem of cooperation is driven by uncertainty about one’s interactions with others in the state of nature. Thus, in this sense, slavery was not cooperative at all: it required the suspension of human rights for a particular group of people such that the rights of all people became more uncertain. Moreover, as Marxian historians and sociologists have long suggested, slavery is an inherently unstable concept that is not conducive to long-term cooperation and is predisposed towards revolution and disorder (see Marx, 1961/1861; Genovese, 1965). Slavery in the United States, for example, led to the outbreak of the American Civil War, the bloodiest conflict in history in
terms of American casualties. Therefore, the example of slavery does not weaken but rather provides support for our conception of moral progress.

With objections aside, I conclude that a standard of moral progress can be defined as follows: morality progresses if and only if it increasingly solves the problem of cooperation. Therefore, I have used one of the four features of the essence of morality to derive a standard of moral progress or metamorality. Next, I will argue that the remaining three features also ought to be included in this metamorality because they are necessary conditions for solving the problem of cooperation.

**Necessary conditions for solving the problem of cooperation**

By logical necessity, it seems that I am unable to conclude that the metamorality or standard of moral progress can include any additional features. Indeed, in its current form, my conception holds that morality progresses *if and only if* it increasingly solves the problem of cooperation. However, I argue that the remaining features of the essence of morality, namely prescriptiveness, perceived universality, and perceived authority over other normative systems, are in fact necessary conditions for morality to perform its function. Therefore, by transitivity, the metamorality or standard of moral progress proposed earlier necessarily includes these three formal features.

**Prescriptiveness**

Morality must be prescriptive in order to solve the problem of cooperation. As shown earlier, prescriptiveness is a necessary feature of any normative system and is thus a necessary
feature of morality. That morality requires prescriptiveness is trivially satisfied and will not be further discussed.

**Perception of universality**

Morality must also be regarded as universalized in order to solve the problem of cooperation. If people do not believe that moral rules are universal, then cooperation remains uncertain. To conceptualize, suppose two individuals in a repeated prisoner’s dilemma are constructing rules to structure their play such that both players cooperate on every turn. Moreover, each player agrees to play TFT: if at any point a player defects, in the next turn, the other player will punish the defecting player. These rules will lead to cooperation only if both players believe them to be universal, i.e. obeyed by both parties. If at any point one player suspects that the rules no longer apply or senses that the other player is unlikely to follow through on the promise of TFT, she might defect to hedge against this uncertainty. Therefore, the absence of universality brings about uncertainty and as a result, cooperation falls apart. This conceptualization illustrates that morality must be regarded as universalized in order to solve the problem of cooperation.

**Perception of overriding authority over other normative systems**

Finally, morality must be regarded as possessing overriding authority over other normative systems in order to solve the problem of cooperation. Let us return to the game theoretical scenario presented earlier, in which both players agree to cooperate on the first turn and play TFT for all subsequent games. If at any point a player suspects that their agreement is overridden by some other agreement or rule, then they might defect in order to hedge against this
uncertainty, again leading to suboptimal outcomes for both players. The common thread of logic between the two situations in this paragraph and the one above is that if uncertainty increases, cooperation is more likely to fail. If rules are neither regarded as universal nor regarded as possessing overriding authority over other normative systems, this increases overall uncertainty. Therefore, the perception of universality and overriding authority of rules are required for morality to solve the problem of cooperation.

Earlier, I concluded that morality progresses if and only if it increasingly solves the problem of cooperation. In Part Two, I demonstrated that solving the problem of cooperation requires morality to be prescriptive, regarded as universalized, and regarded as possessing overriding authority over other normative systems. Therefore, by combining these premises, I update my conception of moral progress to the following: morality progresses if and only if increasingly solves the problem of cooperation and exhibits prescriptiveness, perceived universality, and perceived overriding authority over other normative systems. Therefore, I arrive at the teleological-formalist conception I originally set out to prove: that morality progresses if and only if it increasingly exhibits the formal elements of morality. Phrased differently: morality progresses as it tends closer towards its essence.

In this part, I examined various conceptions of moral progress and argued that each conception is reducible to the teleological-formalist conception of morality, which holds that morality progresses as it increasingly exhibits its essence or formal features.

In the next and final part, I will conclude by considering the implications and practical uses of this conception on moral change.
PART THREE: IMPLICATIONS

In this final part, I will briefly consider implications and practical uses of the teleological-formalist conception of moral progress. In my introduction, I raised the problem that despite extensive research on the evolution of morality, there exists no standard with which to evaluate or normatively relate to moral change. Therefore, I will argue that the teleological-formalist conception of moral progress offers at least two solutions to this problem. First, it allows philosophers to make metamoral judgments about whether moral change is progressive or regressive. Second, the teleological-formalist conception allows philosophers to steer moral change towards progress and away from regress.

Metamoral judgments: evaluating moral change

Before proceeding, it is important to consider whether or not it is possible to detect moral change. On the one hand, detecting moral change can be difficult. First, as suggested earlier, moral change occurs at a rapid rate in evolutionary terms. Due to variability in genotypes that code for human behavior, as well as the influence of non-biological factors, such as culture, morality can change rapidly and unexpectedly (see Richerson & Boyd, 2005). Second, moral change is not always readily apparent or observable. This partially due to the speed of moral change: because morality can change so quickly, by the time we realize that something in our moral systems is now different, the majority of change has already occurred. There is also inherent vagueness encountered when observing morality. When observing moral change, it is difficult to distinguish between significant change and random aberration. We lack tools such as
statistical significance or confidence levels to verify that we are in fact observing a systematic and nonrandom process of moral change.

On the other hand, detecting moral change is possible, even in light of the previous problems. While change may be rapid, we can nonetheless accumulate observations about previous change to inform predictions or projections about the future. For example, scholars who study value change use cross-cultural, longitudinal survey data to extrapolate and predict future data trends (e.g., Halpern, 2001). Adapting this research method to moral change is difficult but not impossible. For example, many studies in the field of business ethics and psychology have used cross-cultural surveys to study moral judgments in business professionals (e.g., Singhapakdi, Vitell, & Leelakultanit, 1994; Tsui, 1996; Davis, Johnson, & Ohmer, 1998). Detecting moral change, therefore, is challenging but possible.

The teleological-formalist conception can be used in a variety of ways to evaluate moral change. For example, empirical scientists can use cross-cultural, longitudinal survey designs to gather information about people’s moral judgments and intuitions. Philosophers can then code the responses such that each is categorized according to different kinds of morality. Once a sufficient number of responses is collected, philosophers can review trends in morality and use the teleological-formalist conception to evaluate whether the trends are morally progressive or regressive.

More generally, the teleological-formalist conception can be used to structure the debate on moral change. Currently, philosophical discussion on moral change is largely unproductive: philosophers appear most concerned with whether or not evolutionary findings can be used to
resolve the debate between moral realists and moral relativists (see generally Boniolo & de Anna, 2006). Philosophers seem reluctant to engage with metamoral questions about whether or not moral change is good or bad. However, these are the kinds of questions that philosophers ought to be asking. If morality changes, then we should be interested in whether or not the moral dispositions of society are progressing or regressing. These answers can be used to inform policymakers of what ultimate metamoral ends our society ought to pursue. The teleological-formalist conception can help steer this debate towards these policy-relevant questions. For example, as contemporary morality increasingly exhibits postmaterialist values, philosophers may want to consider: “Does this change allow morality to better serve its function?” (see Inglehart and Baker, 2000). Fruitful philosophical debate on moral change, therefore, can be achieved by using the teleological-formalist conception to raise questions and guide philosophical inquiry.

**Steering moral progress using the teleological-formalist conception**

The teleological-formalist conception also allows philosophers and policymakers more generally to steer moral change towards progressive outcomes. According to the conception, moral change ought to aim towards a final morality, in which the function of morality is fulfilled and the problem of cooperation is solved. Philosophers can thus steer philosophical debate towards values and moralities that aim towards this goal. As these values are emphasized, they flow to policymakers, who can then create policy to effect moral progress. For example, if philosophers find that the recent change towards postmaterialist morality is a good thing, then policymakers can create laws that nudge society towards greater fulfillment of postmaterialist
values, such as legalization of gay marriage. Therefore, if used correctly, the teleological-formalist conception can lead not only to moral progress but social progress more generally.

In this part, I briefly considered implications and practical uses of the teleological-formalist conception of moral progress. I argued that this metamorality can be useful in at least two ways: allowing philosophers to evaluate moral change and empowering philosophers and policymakers more generally to bring about moral progress. In the conclusion, I will briefly summarize the argument made in this paper.
CONCLUSION

In this paper, I derived a metamorality or standard of moral progress with which to evaluate moral change, namely the teleological-formalist conception of moral progress.

In Part One, I induced the essence of morality from various conceptions of morality. First, I presented and criticized Frankena’s (1966), instead offering my own set of features that I believe compose the essence of morality. Furthermore, I argued that the function of morality is to solve the problem of cooperation.

In Part Two, I derived a conception of moral progress, or “metamorality,” with which to evaluate moral change. In particular, I examined various conceptions of moral progress and proposed the teleological-formalist conception of morality, which holds that morality progresses if and only if it increasing exhibits the essence or formal features of morality.

In Part Three, I concluded by considering the implications and practical uses of this proposed conception. In particular, I briefly described how philosophers and policymakers more generally could use the teleological-formalist conception to evaluate moral change and steer moral progress.

I end with an analogy. In the past century, scientific advances have been remarkable. However, they have also been deeply controversial. For example, the discovery of gestational surrogacy aroused intense moral and religious debates that continue today. Moreover, with nearly every controversial scientific discovery thus far, philosophers have been happy to participate in these debates and engage with profound and challenging questions.
Moral change is the discovery of new values, principles, rules, or systems for guiding human conduct. Just as scientific discoveries can be deeply controversial, so too can moral change raise puzzling questions. Moreover, just as philosophers are eager to engage with puzzling normative questions arising from scientific discoveries, so too should philosophers be eager to engage with normative questions arising from moral discoveries. I have attempted to engage with some of these questions, namely the question of how we should normatively relate to moral change. I do not claim that the teleological-formalist conception provides a definitive answer. Nonetheless, my hope is that I have at least provided some guidance as well as inspiration for other scholars to join me in taking on these difficult questions.
BIBLIOGRAPHY


Kar, R. B. (2012). The two face of morality: How evolutionary theory can both vindicate and debunk morality (with a special nod to the growing importance of law). In J. E. Fleming & S. Levinson (Eds.), *Evolution and morality: NOMOS LII* (pp. 31-99). New York: New York University Press.


