

# **Moral Fictionalism and Moral Reasons**

by

Patrick Clipsham

A thesis  
presented to the University of Waterloo  
in fulfilment of the  
thesis requirement for the degree of  
Master of Arts  
in  
Philosophy

Waterloo, Ontario, Canada, 2008

©Patrick Clipsham 2008

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

## Abstract

One major problem with moral discourse is that we tend to treat moral utterances as if they represent propositions. But complex metaphysical problems arise when we try to describe the nature of the moral facts that correspond to these propositions. If moral facts do not exist, how can moralizers justify engagement in moral practice? One possibility is abolitionism; abandoning morality and growing out of our old habits. Another option that has been suggested is that morality be preserved as a *useful fiction*. Moral fictionalists propose that moralizers come to understand their moral beliefs as fictive precommitments that are instrumentally valuable. In this essay, I argue that this type of instrumentalist justification does not allow moralizers to have genuinely *moral* reasons for acting in accordance with their precommitments. The legislative function of morality and the concept of moral personhood cannot be supported by metaethical theories that only provide instrumental reasons for adopting moral discourse. Ironically, this implies that an instrumentalist moral society would not be able to preserve as many useful moral concepts as would a non-instrumentalist moral society. Since the fictionalists' own criterion demands that they endorse the *most* instrumentally valuable metaethical theory, they cannot persuasively argue that their own metaethical alternative is viable. Fictionalists should, I argue, prefer a non-instrumentalist theory that preserves more of the functions of moral discourse, such as quasi-realism. The argumentative strategy employed by fictionalists is therefore self-undermining.

## **Acknowledgements**

I'd like to thank everyone in the Philosophy Department at the University of Waterloo for their advice and support. I would especially like to thank Professor Patricia Marino, Professor Paul Thagard, and Professor Dave DeVidi for their comments on this essay.

*For Steph*

## Table of Contents

Introduction	1
I: Metaethical Instrumentalism	4
II: Moral Fictionalism as Instrumentalism	7
III: Moral Personhood and Instrumental Reasons	13
IV: In-groups, Out-groups and Revolutionary Fictionalism	21
V: Moral Legislation and Psychopathy	30
VI: Moral Imputation and Revolutionary Fictionalism	42
VII: The Economy of Metaethical Instrumentalism: Quasi-Realism over Moral Fictionalism	53
Conclusion	67
References	68

## Introduction

When we employ moral discourse, we tend to act as if it is representational. Moral predicates are treated as if they identify substantive properties and moral sentences are generally taken to express propositions. Our first-order moral discourse is riddled with these metaphysical presuppositions. When we make moral claims or assert moral commitments, we treat these commitments as if they are somehow *objectively* true. This common aspect of our moral experience tempts individuals into *realistic* explanations of morality and metaethical discourse. This realistic practice interprets moral propositions as being *literally* true or false.<sup>1</sup>

The problem with realistic metaethical discourse is that it falls prey to a number of complex metaphysical problems. Some *anti-realists* find fault in the realists' claim that moral utterances are genuinely assertoric. Other anti-realists conclude that common moral practice is wrong to assume the existence of any kind of moral fact. Philosophers in this vein of thought, usually referred to as *error theorists* suggest that the very foundations of moral discourse are rooted in erroneous assumptions.<sup>2</sup> But if we were to affirm the error-theorists' analysis, where would that leave moralizers? Should they continue to engage in moral practice, or try to rid themselves of this misguided habit? One suggestion that has been put forth is that realistic moral practice should be preserved, not because it is *correct* to make realistic presuppositions about morality, but rather on the basis of morality being a *useful fiction*.

*Moral fictionalism* claims that moralizers could correct their erroneous assumptions while preserving some of the benefits of moral practice. Fictionalists suggest

---

<sup>1</sup> See Sayre-McCord (1988) p. 6

<sup>2</sup> See Mackie (1977).

that they do so by treating their moral commitments as they might treat facts about a fictitious universe (eg, ‘Sherlock Holmes lives on Baker Street’ is a fact within the fictional world of Sir Arthur Conan Doyle’s literature). One might, then, be able to assert moral claims, act in accordance with those claims and require others to do the same while believing that the entire domain of moral discourse is based on fictitious precommitments. This fictitious attitude might be able to preserve some important features of moral discourse without relying on questionable metaphysical assumptions.

Moral fictionalism relies on an argumentative strategy I refer to as *metaethical instrumentalism*; it attempts to provide reasons for engaging in moral practice that focus only on the instrumental value of doing so for the individuals involved. My primary concern in this essay is to demonstrate that moral fictionalism cannot preserve many of the important functions of moral discourse; fictionalism undermines the concept of *moral personhood* and threatens the peculiarly *legislative* function of moral language. Reliance on metaethical instrumentalism threatens more instrumental benefits of moral practice than it stands to preserve. Since moral fictionalism cites instrumental advantages as the primary reasons for engaging in moral practice, the theory’s inability to support a number of useful moral concepts raises some serious questions about the viability of the fictionalist project.

In the first two sections of this paper I outline the relevant features of metaethical instrumentalism and discuss one prominent theorist who employs this strategy, Richard Joyce. The following four sections are devoted to outlining important features of our moral practice that could not be captured by a moral fictionalist society. In the final section of this paper, I argue that the moral fictionalist’s instrumentalism is self-



undermining. Since the fictionalist is only concerned with the instrumental benefits of a moral system, she should prefer that we correct our erroneous moral assumptions in the most economical way. The fictionalist would, by her own argumentative strategy, have several reasons to prefer that moral fictionalism be rejected in favour of other non-instrumentalist metaethical theories, namely a prescriptive quasi-realist theory.

## I: Metaethical Instrumentalism

My first order of business is to clarify how I plan to use the term ‘instrumental metaethic’ and my motivations for doing so. The term ‘instrumentalism’, as it pertains to metaethical discourse, is sometimes used to indicate various forms of non-cognitivism. These approaches are distinguished by their attempts to alleviate complications in explanations of moral utterances by offering an interpretation of them as being expressive rather than assertoric.<sup>3</sup> I do not wish to directly argue for or against any particular theoretical taxonomy of moral language. This essay is not directly concerned with whether or not moral language is genuinely cognitive or whether it is expressive. My own use of the term ‘instrumental’ has to do with the *kinds of reasons* a metaethic can provide for individuals to act in accordance with moral principles. The term ‘instrumentalist’ is ideal for my purposes because it identifies the central component of these theories that is at issue. An instrumentalist theory is one that justifies moral practice in terms of a cost-benefit calculation for the individuals engaged in the practice.<sup>4</sup>

The instrumentalist approach is characterized by the attempt to categorize and qualify the groundings of moral discourse in the same way as one might explain any other behaviour. This perspective depicts moralizers from an anthropological, sociological or psychological standpoint. The assumption that underpins this metaethical attitude is that moral cognitive processes and linguistic practices can be accounted for as could any other behaviour in any other type of organism.

---

<sup>3</sup> See Sayre-McCord (1988)

<sup>4</sup> My use of the term instrumental is also indebted to Yablo’s (2001) articulation of instrumental fictionalism.

A colony of ants may cooperate to the extent that the welfare of the individual may often be secondary to the benefit of the entire colony. A myrmecologist studying the colony could tell an explanatory story of how this behaviour developed and what types of neural mechanisms favour the propagation of this behaviour. Colonies wherein the individual has a predisposition towards allowing himself to be sacrificed for the sake of the queen best facilitate the production of future generations. It is because of this fact that cooperative behaviour is 'programmed' into each individual ant. Additionally, this type of account could act as a reason why any other group *should* cooperate; the instrumentality of cooperation vindicates engagement in that particular type of behaviour.

Note that instrumental metaethical strategies differ from social contract theories. A social contract theory is geared towards the vindication of a particular *type* of moral and political system; monarchies, liberal democracies and other types of government institutions have been supported on the basis of a theoretical social contract that maximizes the benefits for all involved.<sup>5</sup> Instrumental metaethical strategies do not provide reasons for engaging in any particular type of moral system, but rather deal with the reasons an individual might have to engage in any moral practice at all.

These instrumentalist explanations characterize all of moral practice as a means of accounting for the problems of collective action, akrasia and cooperation in human societies<sup>6</sup>; in short, they offer practical non-moral reasons for acting on the basis of moral precommitments and engaging in moral discourse. From this perspective, the grounding of moral practice would have to come about in the same way that the behaviour of the ant colony would be vindicated; a cost-benefit analysis, justifying the behaviour in terms of

---

<sup>5</sup> See Rawls (1999a) and Hobbes (1994)

<sup>6</sup> There will be a more detailed explanation of how morality is taken to facilitate these benefits in later sections (See Joyce 2001, 2005).

biological adaptation, social pressure and the evolution of human society. A metaethical instrumentalist, if asked to explain her reasons for engaging in moral discourse, would respond by saying 'I am being moral because it best serves my long and short-term interests to do so'.

My contention is that relying on these types of instrumental reasons for engaging in moral behaviour undermines one of the primary functions of moral language and threatens our use of certain important and useful normative concepts.

## II: Moral Fictionalism as Instrumentalism

Richard Joyce offers one of the more persuasive accounts of an instrumental metaethic. He employs this strategy to justify a form of revolutionary moral fictionalism. Joyce is concerned with the possible options available for moralizers once they acknowledge the dysfunctional nature of realistic moral practice. Joyce only offers two options to moralizers who find themselves in this position. The first possibility is the abolition of moral discourse; they simply grow out of the habit of making first-order moral assertions. The second option is the acceptance of moral discourse as a useful fiction.

It is important to note the two traditional ways a fictionalist theory can be understood; hermeneutical interpretation and revolutionary prescription.<sup>7</sup> A hermeneutical moral fictionalism would hold that we moralizers actually *do* see moral discourse as a fiction. Revolutionary fictionalisms are prescriptive in that they *recommend* that we start to employ the discourse as is depicted by fictionalist theories. The central claim of the revolutionary moral fictionalist is that morality would still be a useful tool if we *began* to look at it in terms of only its instrumental value and employ it as we would any other fictitious discourse. This would presumably resolve the problems related to the use of realistic moral practice while allowing moralizers to continue to employ some aspects of moral discourse.

My criticisms of moral fictionalism in later chapters are meant to apply to both revolutionary and hermeneutical accounts of instrumentalist metaethics; part of my project is to explain how my criticisms of the hermeneutical project entail certain

---

<sup>7</sup> I use the terms as they are defined in Burgess (1983).

challenges for the revolutionary instrumentalist. Few theorists seriously consider hermeneutical moral fictionalism a viable project.<sup>8</sup> Revolutionary moral fictionalism is, however, sometimes offered as a plausible alternative to realistic moral discourse. Richard Joyce offers one of the more persuasive arguments in favour of revolutionary moral fictionalism.

He defines a plausible fictionalist story as one that gives reasons why we might take “a positive attitude...towards a theory which does not involve believing it”.<sup>9</sup> This positive attitude is probably best described as employing the discourse realistically, despite our knowledge that these realistic presuppositions are grounded in error. In other words, we could be justified in employing a fictionalist reading of a discourse so long as “we gain a benefit from engaging”<sup>10</sup> in the practice of employing that discourse.

Moral fictionalism seeks to offer reasons for individuals to engage in moral discourse while conceding that the objects of said discourse do not actually exist or are erroneously applied by moral practitioners. Joyce suggests that if we were to accept the erroneous foundations of morality, our long and short-term goals would be most adequately served by adopting a form of non-cognitivist moral practice. This would allow moralizers to continue to utter moral expressions without committing themselves to erroneous beliefs. For Joyce to justify the adoption of this non-cognitive fictionalism, he only has to identify one instrumental benefit that his fictionalists would gain that would not be available to abolitionists. He does this by explaining how moral practice helps moralizers combat akrasia.

---

<sup>8</sup> See Joyce (2001, 2005), Kalderon (2005), Nolan et al. (2002)

<sup>9</sup> Ibid, Pp. 204-205

<sup>10</sup> Ibid, p.199

For Joyce, controlling akratic impulses is one of the more important instrumental values of moral practice. Since he presupposes that error-theory about morality is largely successful, his conception of a revolutionary moral fictionalism would depict moralizers in a fictionalist society as being little more than sensible knaves attempting to best achieve their long-term goals<sup>11</sup>. As knaves, it would best serve our interests to attempt to forge strong trusting relationships with those around us. For most people, being found out as a traitor would hinder their achievement of personal goals. Presumably, however, we sensible knaves will sometimes be tempted to violate any codes of conduct that our confederates have established.

Even if we have Humean (emotional) and Hobbesian (legal or political) reasons to continue to abide by moral rules, it is inevitable that we will eventually be tempted by the possibility of violating them. This type of rational calculation can, however, be dangerous. The more often we entertain knavish thoughts and engage in knavish actions, the greater our chances of being caught violating a moral principle and, thus being excluded from the community of moralizers. Joyce likens the value of moral fictionalism to a mental exercise, or a particular ‘thought’ more so than a system of belief (the distinction between a thought and a belief, for Joyce, being that the former does not facilitate a tendency to dissent).<sup>12</sup> He asks us to consider various mental strategies that would help us motivate ourselves to reach a certain goal for daily exercise. The heuristic he elaborates is the rule ‘do fifty sit-ups every day’. In some conversational contexts we may admit that most days forty sit-ups would suffice, but when we are actually engaging in the activity itself, we tell ourselves that we must do fifty. Thus, we do not truly *believe*

---

<sup>11</sup> Ibid p. 208

<sup>12</sup> Ibid p. 219

that we must do fifty sit-ups every day, but rather employ this thought in such a way that it motivates us to do so. The obligation expressed in the thought is fictitious, but since it serves a relevant purpose, we continue to employ it.

This type of mental process could also help individuals overcome other akratic drives, such as the temptation to commit a moral wrong at the risk of being caught. Any reasonable instrumentalist will realize that in some circumstances, the chances of being caught in a moral infraction are fairly low. But the more often we rationalize such infractions, the higher our probabilities of being discovered. Precommitting oneself to moral discourse as a functional facilitates avoiding being discovered as a knave; this motivation is comparable to Odysseus tying himself to the mast of his boat or Richard Joyce smearing gravy on his cheesecake to ensure that he will not finish it.<sup>13</sup>

My argument is not geared towards refuting Joyce's claims regarding the value of moral discourse in combating akrasia. I will grant that he has convincingly argued for one instrumental value that would be preserved by a moral fictionalist stance that would be lost by moral abolitionists. But, though he does an admirable job enumerating *some* instrumental reasons for a move towards moral fictionalism, he does not consider what would be lost in such a revolutionary shift from realistic to fictionalist moral practice.

Joyce's fictionalism employs the theoretical tendency I described above as metaethical instrumentalism; the attempt to identify the reasons for engaging in moral practices as being dependent only on the instrumental value of doing so. For Joyce, the only reason to prefer fictionalism over abolitionism is because of the practical benefits of doing so; adopting moral fictionalism allows us to continue to use an instrumentally valuable mental exercise. But, insofar as his fictionalism is supported only by economical

---

<sup>13</sup> An example of one of Joyce's own practices for overcoming akrasia. See Joyce (2005).



considerations regarding the benefits of continuing to engage in moral practice, Joyce and other instrumental metaethicists should also consider the possible drawbacks of participating in their proposed moral systems. If there are significant disadvantages that would follow from the adoption of an instrumentalist metaethical strategy, the moral fictionalist could only endorse a revolutionary shift towards fictionalism if there were no non-instrumentalist metaethical alternatives.

My contention is that there are alternative theories that fare better than metaethical instrumentalism in terms of moral fictionalists' own standards by which they defend their position. In other words, there are alternative theories that successfully liberate moral practice from mistaken realistic presuppositions, while preserving *even more* of moral discourse's instrumentally-valuable functions. Revolutionary fictionalism is defended by Joyce because it preserves more instrumentally-valuable moral practices than does abolitionism. Joyce's argument thereby relies on the rule that we should endorse whichever metaethical theory best preserves these instrumentally-valuable moral functions.

In the following sections of this paper, I argue that a fictionalist moral society would fail to uphold a number of useful, instrumentally-valuable moral concepts and functions. In particular, I argue that a move towards moral fictionalism would entail the abandonment of the concept of a moral person, a move from an out-group to an in-group moral system and a failure to uphold the legislative function of moral discourse. Such a society's inability to preserve these aspects of our moral lives suggests that the fictionalist alternative might not be as attractive as Joyce and other theorists would lead us to believe.

In the final section of this essay, I argue that quasi-realism would account for anti-realist concerns about moral discourse while preserving more instrumental functions than would moral fictionalism. Since Joyce and other fictionalists endorse whichever theory best preserves the instrumental functionality of moral practice, fictionalists should thereby prefer the success of a revolutionary quasi-realism over the success of their own project. In order to persuasively argue for their conclusion, revolutionary moral fictionalists would have to provide some reasons to doubt the success of a quasi-realist project. Pending such a rejection, the prospects for revolutionary moral fictionalism are dim. The onus now falls to my shoulders, however, to demonstrate the instrumentally valuable moral practices that would be lost in a shift towards moral fictionalism.

### III: Moral Personhood and Instrumental Reasons

An important feature of our moral life is the belief that moral persons (those individuals to whom we attribute moral praise and blame) are capable of giving good reasons for adopting, supporting and propagating their moral commitments. Any individual who cannot offer reasons for supporting the moral behaviour she engages in is blindly participating in an institution without concern for whether or not that institution is justifiable. We only hold an individual *morally* responsible for her actions if she has the cognitive capacity to judge whether or not participation in a moral system is a good thing to do. Unless an individual is capable and willing to show *concern* about what types of reasons underpin her moral evaluations, this individual is not considered a fully-functioning moralizer.

The ability to question the groundings of our moral practice is an important feature of those individuals deserving moral responsibility. Lacking this capacity, an individual begins to more closely resemble a wanton; someone who uncaringly follows whichever principles, beliefs and desires drive her.<sup>14</sup> The distinction between moralizers and wantons (classically characterized as those individuals who do not care which of their desires become effective) has lately been considered one of the fundamental capacities that contribute to the attribution of moral status.<sup>15</sup>

But if moralizers are believed to be able to care about which of their motivations drive them, it follows that they should also care about what grounds the practice of valuation that facilitates this caring. A truly moral agent, in other words, should care

---

<sup>14</sup> See Frankfurt (1971)

<sup>15</sup> See Frankfurt (1971), Dennett (1981).

about whether or not she has *good reasons* to endorse any of the particular states of affairs that she does, in fact, endorse.

This not to say that genuine moral agents must always question every single motivational principle they adopt. Most people accept rules like ‘try to be healthy’ without every directly questioning why it is that they should desire good health. But if we were to question moral persons regarding why they value good health, we assume that these individuals would be able to provide some reasons that justify this commitment. People often cite a fear of death, a desire to maintain a certain weight or other motivating factors. We also expect that these individuals would be able to provide reasons for valuing long life or a thinner waistline. The important point is that moral persons are assumed to be capable of providing such reasons for all of their evaluative commitments, including their moral beliefs.

An important aspect of our attributions of moral status is thereby the presupposition that an individual would be able to formulate reasons why she values what she values. Consider hypothetical individuals that value poor health, but have no reasons to offer regarding why they value poor states of health. They stubbornly assert their normative commitments and thus smoke, drink and eat gluttonously. Without being able to consider the reasons they have for valuing poor health, there is no conceivable means by which these individuals could revise their normative commitments. Lacking the capacity for such deliberation, an individual begins to more closely resemble someone who has been brainwashed or otherwise indoctrinated into a belief system against her will.

My argument can be best outlined in terms of an example. Consider two individuals; one is a wanton and the other purports to be a rational moral agent. Both are offered two very similar research jobs. One is a highly paid position at a company that studies and produces biological weapons.<sup>16</sup> The other is a lower income job that audits large companies on their environmental impact. The wanton has no concern regarding which of her desires win out, even though she is driven equally by her desire to acquire money and her concern for the welfare of the environment. She chooses the higher paying job and, when asked why, responds that she did not care which desire moved her.

The self-identified rational moralizer, however, feels compelled by her moral sentiments to favour the environmentally-friendly research position. As such, she desires that her interest in preserving the environment be effective and that she take the lower-income position. But, when asked why she favours environmental research to the research on biological warfare, she has no reasons to offer. Preservation of the environment is simply a morally valuable principle to her and she is incapable of offering reasons why concern for the environment is morally important.

Is the so-called rational moralizer really any better than the wanton? If she does not care about the kinds of reasons that underpin her moral appraisals, her moral commitment in favour of environmental research is arbitrary. She acknowledges that her moral commitments are entirely contingent. Additionally, if she were to be given reasons regarding *why* her moral beliefs were mistaken, her lack of concern for the veracity of her moral beliefs would entail that such arguments would fail to change her mind or influence her decisions. This stubborn individual, unwilling or incapable of revising and

---

<sup>16</sup> The style of this example was inspired by one offered by Williams (1973), but the scope and purpose of the example are distinct.

questioning her moral commitments is not much closer to being a moral person than is the wanton. They do differ, in that the wanton fails to pay any attention to metaethical questions, whereas the so-called rational moralizer is just wholly unconcerned about the answers to such questions. But neither individual is capable of successfully justifying or revising her moral commitments. Thus, the same moral complaint that applies to the wanton applies just as much to the so-called moralizer. Lacking good reasons for endorsing a particular moral system, an individual is not considered to be a full-fledged, reliable moralizer.

Moralizers want to have good reasons for acting morally; being unable to offer such good reasons should lead a rational moralizer to revise her moral commitments. Some theorists have contended that reflection on one's own commitments might not always diminish the strength of said commitments.<sup>17</sup> I am not claiming that reflection on the erroneous foundations of moral practice eliminates all attachment to a moral commitment. All I am arguing is that we assume that other moralizers have good reasons for the adoption of their moral systems; this assumption is part of our common moral practice. But moralizers should not only care that they have *some* reasons that underpin their moral commitments, but should also be concerned with having *more* than just instrumental reasons. This is because only non-instrumental reasons can lend themselves to the vindication of the use of the concept of a moral person.

Consider individuals who only accept instrumental reasons for engaging in moral practices; how would they conceive of their roles as moral persons? All individuals that engage in moral behaviour must have reasons for doing so. For the sake of clarity, keep in mind that the type of instrumental reasons under discussion here are similar to those

---

<sup>17</sup> See Blackburn 1993a.

cited by Joyce regarding akrasia. An instrumental reason for engaging in morality is somewhat egoistic. For example, Joyce depicts moralizers as engaging in moral discourse for the sake of preserving a mental practice that leads to long-term benefits. For an individual to have an instrumental reason to engage in moral discourse, it must make sense for *her* to engage in moral behaviour based on her long-term goals and the consequences of participation in moral practice. The most intuitive and common example of this type of reason would rely on a cost-benefit analysis demonstrating that having her peers trust her is more beneficial than having them distrust her. Recall once again the argument Joyce offers concerning akrasia. Sustaining a mental exercise that combats akrasia helps the individual in question who employs the exercise; it helps her avoid being caught as a violator of moral principles. This depiction of moralizers as individuals who engage in moral practice only because of the instrumental value of doing so implies that all moral behaviour is actually egoistic.

This is a largely counter-intuitive depiction of how people give reasons for their moral behaviour. Truly moral behaviour requires acting on the basis of accepted moral principles and beliefs; this point may be contested by some non-cognitivists, but this conceptual architecture is central to realistic moral practice. An individual who only endorses moral behaviour out of instrumentalist concerns, however, has the capacity to recognize, in some contexts, that *only* the instrumental reasons govern her assent to any moral proposition. She thereby assents to being moved in a way that is consistent with her *believing* a moral principle to hold, while at least tacitly acknowledging that the moral principle is not what actually moves her; egoistic concerns are what *actually* motivate her moral behaviour. This individual exhibits a lack of concern for what types

of assertions she assents to and what beliefs motivate her actions; she commits herself to whatever principles help her gain an advantage. This individual does not possess the characteristics that we generally deem valuable in moral persons. Her perspective towards her own beliefs more closely resemble that of a wanton than a moral person.

This is because we moralizers simply *do not* see moral behaviour as *only* possessing instrumental value and socially fortuitous consequences. If we did see our moral practice in this light, there would only be sensible knaves; individuals paying lip service to a legislative system, while privately endorsing other standards of value (standards of value centering on predominantly egoistic concerns). This type of practice is not what we characterize as genuinely *moral* behaviour. It is far from obvious that obeying the law *only* out of a fear of being prosecuted constitutes genuinely moral behaviour. This intuition is clearly formulated in Kantian moral philosophy, in which one's reasons for acting in accordance with moral principles are the only measure of the moral worth of any act. This same theme, however, has been invoked in a staggering amount of moral and metaethical theory.<sup>18</sup> This characteristic of moral practice is a widespread, well-supported intuition.

There is evidence available for this claim via the consideration of some of our most commonly invoked moral intuitions. If an individual were to find a loophole in the legal system that rendered a morally deplorable action legally permissible, would this act then cease to be morally reprehensible? Similarly, if all liberal democratic governments were suddenly replaced by foreign imperialists who condone the complete subjugation of a certain race, the moral repugnance of enslaving that race would not diminish. In section four of this essay, I recount evidence from studies performed by James Blair that suggest

---

<sup>18</sup> See Blackburn 1993a.



most moralizers understand a relevant distinction between conventional codes of conduct and moral principles.<sup>19</sup> The distinction between conventional codes of conduct and moral principles is meant to be brought out by the example about a foreign imperialist nation. The goal of this example is to contribute to my argument that realistic moralizers see themselves as *more* than just individuals whose behaviour is regulated by an attempt to gain the greatest advantage over others. Genuine moralizers are not *just* sensible knaves; we do not only see ourselves as paying lip-service to a system of rules because it pays for us to do so, as would a psychopath, but we believe that the system of rules we have adopted is *the right one*.<sup>20</sup> Even if it ceased to be expedient to continue to act in accordance with the traditional standards of conduct, as in the example above, empirical evidence suggests that most moralizers would continue to identify their own moral beliefs as having greater legislative authority than these conventional principles.

In this section, I argued that there is strong evidence to suggest that most moralizers *do* identify moral systems as being grounded on more than just instrumental reasons. Contemporary philosophical literature on the conception of moral personhood suggests that the ability to reflect metaethically on *one's own* moral practice, without reducing the reasons for engaging in this practice to the purely instrumental level, is an important component of what it means to be a moralizer. If moralizers were only concerned with the instrumentality of our moral practice, a global shift in conventional codes of conduct would immediately and radically effect their moral reactions to various situations; this conclusion is, however, contrary to empirical evidence about how we understand and employ moral concepts. The following section of this chapter is

---

<sup>19</sup> Blair et al (1995)

<sup>20</sup> For a detailed account of this aspect of realistic moral practice, see Dworkin (1996), p. 92.

concerned with demonstrating why this poses a significant problem for revolutionary moral fictionalism. Since revolutionary fictionalism suggests that we start to understand our moral practices in a fictional context, this society would fall to prey to the arguments elaborated in this section. Moral fictionalists would not be able to coherently employ the concept of a moral person. I argue that the inability to uphold this concept leads to unfavourable consequences for any society that adopts moral fictionalism.

#### **IV: In-groups, Out-groups and Revolutionary Fictionalism**

Consider again the individual employing an instrumentalist metaethical strategy to account for the groundings and nature of moral discourse and the moral behaviour of a particular group. She could only locate herself within the studied group of moralizers if she accounts for her own moral practices as being *only* instrumentally valuable. In order to understand what may be unfavourable about this state of affairs, recall the analogy I proposed earlier comparing human moral societies to an ant colony. The behaviour of the ants is justified in terms of the overall benefit for the entire colony of insects; the ants cannot survive independently of the colony, so it makes sense for each individual to work towards the preservation and propagation of the ants within the colony.

It is obvious that this analogy does not carry over directly to the case of humans. The explanation for the cooperation of the ants is largely evolutionary; the only reason the ants' ancestors survived is because they exhibited these kinds of behaviours. As such, the behaviour facilitates reproduction of genetic material but fails to do what is really best for the individual.

But with a bit of creative anthropomorphism, we can begin to see the difference between a purely instrumental social institution, like that of the ants, and a moral institution that is thought to be valuable for more than instrumental reasons. Suppose that the ants' cognitive capabilities were in fact similar to those of a human; they can engage in higher-order reflective thought, complex language, they can philosophize, etc. The only difference is that for most of their lives, they *don't* engage in these activities. They continue to act according to the stringent 'ant code of conduct', which instructs young

ants to work inside and protect the colony, and older ants to go searching for food to sustain the rest of their brethren. But, when pressed on their behaviour (say, by a human myrmecologist who has learned their language), the ants are capable of discussing their behaviours at an abstract, ‘metacolonial’ level. When asked why they continue to bring all the food they find back to the colony rather than partaking in it themselves, the ant may at first simply cite the rules of the colony.

But, if pressed to his most critical reflective context<sup>21</sup>, the ant would be forced to concede that, really, he would likely be better off if he consumed the tastiest morsels himself or even if he hid somewhere near the colony instead of risking long excursions to find food. The ant, however, does not succumb to these knavish suggestions and thus justifies his behaviour as best suiting his long-term goals. Without the trust and support of his fellow ants he would have no protection from the elements, would be susceptible to attacks from predators and would have no consistent source of food.

In his most critical context, the ant justifies his cooperative behaviour and the social institutions that support this behaviour in terms of the instrumentality of such practice. He *acts* as would the unthinking ant, obeying his evolutionary formula, because he, when pressed, understands that obeying these rules best fulfills his long-term goals.

Let us now return from this anthropomorphic flight of fancy to the case of human moralizers. Presumably, it would be possible for people to engage in moral behaviour six days a week but then, on figurative Sundays, concede that the only reason they do so is

---

<sup>21</sup> This is how Joyce characterizes the human ability to negate the claims of a discourse in one context and yet continue to use the discourse to some end in another context (Joyce 2001, pp. 192-193). For Joyce, critical thought is based on analyzing other kinds of thoughts; it is only in such a critical context that we could question the nature or veracity of moral thoughts. Nolan et al. characterize the entering of this context in terms of the identification of certain bridge principles; principles that explain when and how we move between a base discourse and a fictive discourse (Nolan et al, p. 6)

because conforming to that behaviour best facilitates the achievement of their long-term goals. Like the ant in the example, these individuals' most critical contexts admit the erroneous assumptions that are latent within moral practice and offer an alternative reason for continuing to engage in the behaviour. Their account of morality thereby resembles a cost-benefit analysis; the type of explanation of the behaviour that would be performed by an anthropologist, sociologist or psychologist seeking to account for and offer an instrumental justification of the behaviour. This is, presumably, a rough characterization of how a moralizer would reason about morality if she were to fully accept a purely instrumentalist account of her moral practice.

When placed into their most critical context, these individuals would only be able to offer instrumental reasons for engaging in moral discourse; their reasons for being moral boil down to what, overall, is best for them. It is troubling to think what kind of moral theories could be justifiably supported only in terms of the instrumentality of the theory itself. Lacking any non-instrumental grounding for moral judgement, we lose the ability to discern qualitative differences between particular kinds of instrumental reasons. For example, an individual who has been raised by violent racists might justify participation in her own moral system (in which violent assaults on racial minorities are required and praised) based on the fact that it best serves her personal utility to remain close to the tight-knit, well-armed group to which she belongs. Suppose, as well, that she knows that if she were to attempt to leave the group or question their motives, she would be killed. In this instance, it is instrumentally valuable for her to continue to endorse the moral system that permits violent racially-motivated assaults, just as it was instrumentally valuable for the ant to continue to participate in the activities of the colony.

The metaethical instrumentalist offers no means by which a moralizer could justifiably judge various types and tokens of instrumental reasons; there is no relevant way to discern between an Allied soldier who believes in the long-term value of liberal democracy and a Nazi soldier who (because of his own racist commitments) believes that the Third Reich will establish a utopian society. Both can justify their behaviour in terms of the instrumentality of acting as they do, but most moralizers would demand that the Nazi be held responsible for his faulty moral beliefs. There is no morally relevant way for these instrumentalists to discern between instrumental reasons that are based on a love of democracy and instrumental reasons that are based on racist commitments. Part of our current conception of moral persons is that they are capable of discerning a difference between these two types of reasons, but metaethical instrumentalism has no means by which to justify this conception of persons.

If we adopt metaethical instrumentalism as a valid explanatory tool and thus admit that our only reasons for being moral come from a desire to preserve our advantage, there is no way to sustain a pertinent difference between in-group and out-group moral systems. These two terms are generally employed to indicate the divergent attitudes held towards people who are identified as being within or without some relevant social class. I use the terms to refer to the way a moral system sets down standards regarding individuals from different cultures and communities.

An in-group morality is based on solidarity amongst a group of individuals; there is one standard of conduct that is to be applied to those individuals within the group, and another standard applicable to those who are outside the group. An out-group morality does not sustain diverse standards of conduct for individuals within and without a

particular community. Out-group moral systems recognize the importance of universality and impartiality in their moral principles and do not allow for divergent standards of conduct based only on community-membership. It is clear that the way we currently perceive and employ moral practice is closer to the out-group than the in-group conception. We do not draw arbitrary lines regarding which individuals deserve what kind of treatment, but rather base our treatment of individuals on their cognitive capacities, their past actions and other relevant beliefs about them. These designations avoid being arbitrary because the criteria they make central to the designation of moral status, such as cognitive capacity and the structure of the will, are directly relevant to the individual's personal capacity to involve themselves in a moral community.

The instrumentalist judges a moral system based on how well that particular moral system facilitates the achievement of the long and short-term goals of the community that adopts it. But this deliberative process does not lend itself to the derivation of an out-group moral system, but rather an in-group. As instrumentalists, our only means of judging what type of moral system we should adopt is based on the expected return of doing so. But an in-group system preserves all the advantages of an out-group system (cooperation, trusting confederates, controlling akrasia), while gaining the additional advantages of being able to exploit individuals who are not members of the community. Instrumental metaethical theories cannot be used to justify an out-group moral system.

By recommending that a community should make the conceptual shift to a form of moral fictionalism, the revolutionary fictionalist would have us abandon the conception of moral persons as individuals who care about the types of reasons that

underpin their moral commitments. Since instrumental metaethical theories render moral persons incapable of distinguishing between various types of metaethical justifications, they fail to be able to identify a relevant moral distinction between in-group and out-group systems. The revolutionary fictionalist tacitly endorses a move from out-grouping to in-grouping. Her own standard of judgement that focuses exclusively on the instrumentality of a moral system requires this. But if endorsing moral fictionalism implies endorsing an in-group moral system, we must wonder if this is the kind of conceptual shift that moralizers might be comfortable accepting.

Our moral practice requires of moral agents that they go further in their justifications of their moral commitments than this type of instrumental practical rationalization; we require that moral persons be able to justifiably ground their moral alignment and explain their reasons for sustaining their contribution to a moral system. Part of why children, animals and other individuals generally treated as wantons are not subject to full moral responsibility is because they do not have the capacity to ask questions about the principles that dictate their social behaviour. They are incapable of entering into what adults might consider their 'most critical context' in order to question the justifiability of certain institutions and types of moral practices. For such a wanton individual, one reason for acting morally would be as good as any other; qualitative moral norms about the *types* of reasons that should be included or excluded in metaethical discourse cannot be supported by instrumentalist metaethical deliberation.

It makes sense for us to deny that individuals who are incapable of discerning qualitative moral differences should be bearers of moral responsibility. Part of an out-group moral system is the identification of the traits and characteristics that are relevant



to an attribution of moral personhood. In an in-group system, there is no need to identify the morally relevant traits of persons because moral status is conferred only via membership in the particular community. The exclusion of other cognitively-developed individuals from the in-group moral system is arbitrary and based only on contingent circumstance.

If we begin to base our moral decisions on only instrumentalist rationalizations, we are not only left with an in-group society, but we are also forced to concede that we have no justifiable means by which to determine who should be considered within the community and who should be considered without. We only allow individuals into the community if, in the long run, we would benefit from doing so. Adopting this type of moral system runs contrary to some of our most demanding moral practices; there would, for example, be no consistent way to excuse any member of the community from moral responsibility and blame. Whether or not an individual was cognitively developed, fully rational or otherwise deserving of full moral responsibility is not a relevant question in the in-group society; since the society does not base its attribution of moral praise and blame on reflectivity, critical thought or other capacities that widen the gap between wantons and persons, these attributes cease to play any role in the way that we regard the moral value of the actions of other individuals.

The picture of moral society left with us by the fictionalist is therefore one in which there is no relevant difference between an in-group and an out-group system; in fact, an in-group system is preferable, due to the increased utility for the community of moralizers who adopt such a theoretical stance. Furthermore, the traditional notion of

moral personhood has been abandoned by this society. There is no longer any relevant means by which we may decide which individuals are worthy of moral responsibility and which should be excused from full moral responsibility. Cognitive capabilities, maturity and the capacity for critical thought cease to play any role in how and why we decide to praise and blame various individuals. This moral system thereby loses the capacity to establish universal human rights or to offer a means by which young children, cognitively-deficient adults and non-humans might be justifiably excused from the full burden of moral responsibility. Many of our common moral sentiments and intuitions are thereby threatened by the endorsement of a revolutionary moral fictionalism. The community we are left with bears very little similarity to the moral society we originally sought to preserve. Moral society, for the fictionalist, more closely resembles a colony of ants than a kingdom of ends.

For the moralizers engaged in a fictionalist moral system, the locus of moral personhood shifts from attributions of cognitive, critical capabilities to arbitrarily-decided group membership. But this concept of moral personhood is not only beneficial to us because it helps identify the types of people we should consider morally responsible, but it also helps us to determine whom we can trust and whom we need to keep a close eye on. The instrumentality of this particular moral function is lost in the shift from realistic to fictionalist moral practices; this cost is not weighed out by Joyce or other fictionalists. If their entire reason for supporting moral fictionalism is because of its capacity to preserve some benefits of engagement in moral discourse, theorists like Joyce ought to be troubled by the extent to which important normative practices and the instrumental benefits they confer on moralizers would be lost in a fictionalist society. Instrumental

metaethical theories fail to acknowledge that metaethicists are not impersonal anthropologists, sociologists or psychologists seeking to explain the origin of a specific behaviour, but are better understood as individuals trying to determine the best way to live and the right kinds of reasons for doing so. Asking meta-ethical questions about the nature of moral language and the groundings of moral practice is one essential component of this venture.

## V: Moral Legislation and Psychopathy

Reliance on instrumental reasons for justifying moral behaviour also threatens another common, important aspect of our moral practice. Moral discourse is not only a matter of reporting facts or attitudes about a certain circumstance, but rather is meant to convey *magnetism* and *legislative force*.

Magnetism, first articulated by C.L. Stevenson, is the function of moral assertions to move individuals in the direction suggested by the nature of the moral evaluation.<sup>22</sup> A report that something is morally right is not employed as would a report of non-evaluative fact. Moral assertions are meant to provide direct guidance regarding which actions are permissible and which are required. Magnetism is also a characteristic of any other form of evaluative discourse. If I were to evaluate a certain dish as being ‘tasty’, the implication is that my actions should be directed towards eating the dish. Contrast this type of prescription to the reporting of a non-evaluative fact about a dish; that it contains peas, for example. Depending on the tastes of the person I am addressing, the fact that there are peas in the dish could lend itself to divergent recommendations; some people may be moved towards eating the dish, while others might be moved away from the dish. My assertion that there are peas in the dish is not magnetic independently of the personal goals and tastes of the individual I am addressing. The individual’s evaluative pre-commitment about the status of the taste of peas is what regulates how the individual will be motivated in this situation.

But moral assertions carry an extra functional presupposition that is not present in other evaluative domains of discourse. The magnetism we ascribe to moral assertions is

---

<sup>22</sup> See Stevenson( 1937), p. 16

assumed to hold independently of any individual experiencing the assertion's magnetism first-hand. If I offer the moral prescription that person *X* *ought* to do action *Y*, this prescription is meant to hold independently of whether or not person *X* is personally moved towards performing action *Y*. In other words, moral assertions are distinguished from other evaluative forms of discourse by their *legislative* functions.

When one assents to moral assertions, the presupposition is that one's moral commitments should not only inspire action in oneself or someone with relevantly similar goals and tastes, but that the same moral assertions that move one individual should be universally applicable to *all* individuals. This facet of moral language is what distinguishes morality from other evaluative domains. For non-moral ascriptions of the term 'good', such as 'this is good food', the expectation is only that the assertion will have magnetic implications for those who assent to it. An individual who does not like one of the ingredients of the dish will not be expected to be influenced by another individual's assertion that the dish is good.

For moral assertions, however, subjective motivations and personal assent to the assertion are not the only means by which it could have magnetic influence. The beliefs and principles we tend to endorse as being part of our moral practice prescribe behaviours *despite* the fact that other individuals may not experience the magnetism of our own moral commitments first-hand. If I endorse a particular action as one in which an individual was morally obliged to partake, the assumption implicit in this moral statement is that any individual should have acted similarly, regardless of whatever personal motivations might lead her to do otherwise. Magnetism alone is not enough to explain this particularly moral phenomenon. We do not only require that people be moved by

*their own* moral commitments, but we also demand that they be moved by *our* moral judgements. Our moral practice consists largely of the application of seemingly-objective principles that are capable of trumping any individual's personal motivational influences.

There is, in fact, a good deal of empirical evidence that supports my claim that this legislative magnetism is an essential feature of moral practice that differentiates moral assertions from other forms of evaluative language. The evidence in question is mostly derived from neurological studies on individuals whose moral capacities are somehow limited or deficient. Most of the studies performed on this type of dysfunction centre on the phenomenon of the psychopath. After elaborating some background information about psychopaths, I will explain how some particular studies provide evidence that comprehending the distinction between morality and other types of evaluative practices requires the comprehension of this magnetic-legislative force.

Psychopaths are individuals who are most concisely described as seeming to be without conscience; they do not undergo the same emotional experiences in relation to moral judgements as do normally-developed adults. Psychopaths also demonstrate deficiencies in some forms of emotional learning. They do not, for example, react physiologically to some types of emotional stress that would normally cause sweating, increased heart rate and other symptoms in most non-psychopaths. These individuals are not devoid of all emotion, but *are* devoid of complex emotional reactions that pertain to social learning and indoctrination into a moral system<sup>23</sup>. A good example of the type of dysfunction that manifests in these individuals' behaviour is their inability to differentiate between various types of 'wrong' actions. Whereas most individuals can understand the difference between 'morally' wrong acts (such as murder) and 'conventionally' wrong

---

<sup>23</sup> See Hare, 1999, p. 58

acts (such as talking in class), diagnosed psychopaths have difficulty grasping this distinction<sup>24</sup>.

These individuals, who are incapable of distinguishing moral judgement from other types of evaluations are also incapable of comprehending the peculiarly legislative function of moral magnetism. For them, the only reason to abide by moral principles is the outcome of a cost-benefit analysis that focuses on the instrumental value of assenting and dissenting to various moral claims. By employing this calculus, they determine the extent to which they should assent to moral principles in order to maximize long-term benefits.

Non-psychopathic individuals recognize the distinction between principles of action that are assented to based on the instrumental value of doing so and moral principles that demand conformity regardless of the practical reasons one might have to obey or violate them. There is something about the legislative function of moral prescriptions that psychopaths simply cannot comprehend. They are incapable of understanding that certain actions are required only because of the interests of other people. It is their failure to understand this function that limits their capacity to engage in moral practice and exhibit moral behaviours. But what, exactly, are these individuals missing? The answer to this question can be extrapolated from neurocognitive studies of psychopathic symptoms.

Some of the most noted neurocognitive studies of psychopaths have been performed by James Blair and his colleagues; they have attempted to isolate the cognitive

---

<sup>24</sup> Hare characterizes the manifestation of this symptom as them 'knowing the words without hearing the music', creating the image of an individual attempting to fit in with the crowd, but without being able to offer a consistently convincing performance. (Hare 1999, p. 128). This particular phenomena, however, was identified in clinical studies done by James Blair. (See Blair, 1995).

deficiency that acts as a precursor to the development of psychopathy. He and his colleagues have discovered a genetic disposition that reduces the function of the amygdala in those individuals diagnosed as psychopaths. The amygdala plays a central role in the regulation of emotional responses. Emotional reactions commonly related to moral experience, such as admiration or moral disgust are partially dependent on the normal functioning of this part of the brain.<sup>25</sup>

As was mentioned above, moral cognition is not the only faculty that is disrupted by whatever dysfunction causes psychopathic symptoms. In these individuals, many other normal emotional responses are dulled. Emotional interrupt is a phenomenon that can be observed in most human adults; when exposed to disturbing emotional stimuli, conscious processing tends to slow down. No interrupt occurs for either psychopaths or non-psychopaths when they are exposed to neutral stimuli. The theory is that the emotional responses detract from other cognitive capabilities.<sup>26</sup> The exact mechanism upon which this emotional interrupt works is irrelevant; the important point is that disturbing stimuli do not cause the interrupt for diagnosed psychopaths.

The conclusion to be drawn from this fact is that normal moralizing is somehow correlated with the capacity to experience disturbing emotions. This result becomes even more telling when you consider it in conjunction with studies done on primates in which subjects attempt to alleviate the suffering of their conspecifics, even if this alleviation comes at some inconvenience for themselves.<sup>27</sup> Franz De Waal explains how certain studies have found that many primates respond to the suffering of conspecifics with a form of 'emotional contagion'. Perceiving members of their own species suffering can

---

<sup>25</sup> See Blair et al, (2005) p. 117.

<sup>26</sup> Ibid, p. 119

<sup>27</sup> See De Waal, p. 26.



and will lead primates to attempt to alleviate the suffering, even if doing so involves refraining from eating or engaging in another uncomfortable activity. It could be that we, as social beings, are equipped with strong emotional regulators that control the way young humans perceive the emotional reactions of their relatives and confederates. If we see a human in serious distress, many of us experience strong emotional reactions and attempt to stop the troubling event.

I do not wish to suggest that this emotional tendency is tantamount to any kind of ‘moral sense’. It is clear that this emotional reaction is far from being universally acquired throughout the human species; approximately 1% of all people are diagnostically psychopaths.<sup>28</sup> Additionally, pressing circumstances will sometimes override this emotional reaction. The often-cited examples are the Milgram studies from the 1960s, which demonstrate that most individuals will be willing to torture another subject if the proper social pressure is applied. I would, however, like to re-iterate that this emotional response is not innate, is not universal and is not the *only* cognitive mechanism underpinning our behaviour as moralizers. I am characterizing this emotional response system as a necessary precursor for the acquisition of what we recognize as moral concepts and principles. But even though many of the subjects’ moral emotions failed to stop them from torturing the other individual, most of the people who had gone further in the experiment than they expected subsequently suffered emotional stress.<sup>29</sup> This suggests that these emotional regulators are fairly common and widespread.

The deficiency in psychopaths has nothing to do with the dysfunction of a peculiarly ‘moral’ cognitive structure, but rather a disruption of the emotional processes

---

<sup>28</sup> See Hare, 1991

<sup>29</sup> See Doris, 1998

that would normally facilitate moral learning and the reification or objectification of moral principles. Superficially, this evidence seems to suggest that a point has been scored for classical non-cognitivists. Lacking the emotionally charged attitudes that tend to underpin our moral utterances, no moral experiences occur. If an individual is incapable of normal emotional reactions, this individual is also incapable of moral cognition.

But this, fortunately, is not the end of the story. Other research suggests that psychopathic symptoms are not only caused by amygdala dysfunction and the correlated emotional deficiency, but can also be caused by damage to the ventromedial prefrontal cortex. This part of the prefrontal cortex, like the amygdala, is thought to play a crucial role in the processing of moral emotions. It should not be surprising that damage to this area, much like a genetic predisposition for a dysfunctional amygdala, can cause psychopathic symptoms<sup>30</sup>. Some of the more interesting studies on this cortical area have found odd dissimilarities between individuals suffering this damage early in life and those suffering damage as adults.

The individuals suffering cortical damage later on in life begin to exhibit some symptoms that are typical of psychopathy; they cease to undergo physiological emotional reactions to disturbing or stressful stimuli (like sweating when nervous) and they no longer act like a normal moralizer when tested for emotional interrupt. All things considered, these individuals' emotions are no longer being physiologically effective. They are cut off from them in much the same way as they are cut off from psychopaths for their entire lives.

---

<sup>30</sup> See Raine, et al, p. 205

The same is true for the young individuals suffering the same type of cortical damage. There is, however one very important difference between the adult and the younger subjects. The moral learning, scruples and social skills that had been acquired by the older subjects does not diminish once the damage occurs. The older subjects are still capable of employing moral concepts accurately, of identifying moral wrongs and engaging in practical social reasoning.<sup>31</sup> The subjects were tested with what the neuroscientists refer to as a ‘moral reasoning task’ that was meant to demonstrate the individual’s capacity for abstract social insight.

This older subject actually retains her knowledge of seemingly-objective moral facts and thus continues to be able to distinguish between moral wrongs and what Blair identifies as ‘conventional’ wrongs. A conventional wrong is a violation of a particular social structure that requires uniformity and is justified only in terms of its functionality. Moral wrongs are characterized by genuine concern for the welfare of other individuals and the recognition of these individuals as persons with rights.<sup>32</sup> Individuals are able to tell the difference between conventional and moral wrongs when they are capable of retaining the knowledge that certain actions are wrong *only* because of the impact on other people. The older subjects in the study mentioned above continue to be able to understand moral situations in this context; they understand that they are expected to uphold certain moral principles regardless of any instrumental reasons they might have for doing so. In other words, these individuals maintain the capacity to understand the legislative magnetism of other people’s moral evaluations; they see other people’s moral judgements and commitments as having a hold on them. The younger subjects however,

---

<sup>31</sup> Ibid, p. 206

<sup>32</sup> See Blair (1995)

do not have any grasp on the legislative force of moral assertions. For them, the only reasons to act in accordance with moral principles is because doing so best facilitates the achievement of personal goals. Being isolated from emotion for their entire lives has cut them off, like the amygdala-based psychopath, from all the emotions that act as a precursor for their moral and practical social learning.

This telling result suggests the importance of basic emotional functionality to the first proto-moral human experiences. I am using the term proto-moral to indicate the stages of emotional response and social experimentation that act as a precursor to the acquisition and acceptance of moral concepts and principles. The research suggests that these proto-moralizers would be very much how crude non-cognitivist and emotivist doctrines might depict us. Any moral utterances we can be said to make at this stage of development really would be tantamount to saying 'boo torture' or 'yay charity'. At this stage, the only reality for the evaluative principles is subjective. If an individual in this position were capable of being pressed on the reasons why she endorses charity or chastises torturers, her only response would be that this is the type of conclusion her emotions have pushed her towards. She feels as if these things are wrong, so she expresses these feelings.

While this type of emotive explanation may suffice for a case in which adults are asked to justify a non-moral evaluation (such as a preference or dislike for a particular dish), it does not suffice for the case of the justification of a genuinely moral prescription. The evidence recounted above demonstrates how it is that moralizers come to be able to identify the uniquely legislative function of moral discourse. We begin our moral life as little more than instrumentalists; the reason we act on our proto-moral intuitions is

because it can be uncomfortable for us to do otherwise. This is suggested by Adriane Raine and Yaling Yang. They hypothesize that our current moral experiences developed from emotional regulators that were present in earlier hominids. According to this research, it is highly probably that many of our most widespread moral commitments are “shrouded in a deep evolutionary history where emotions, not cognitions, constituted the driving force of moral action.”<sup>33</sup> The individuals who suffer PFC damage early in life fail to develop the cognitive aspects of moral discourse. The individuals suffering the same damage later in life *do* develop complex moral cognition. This implies that the emotive side of our moral experience must be the first to present itself in children; their conscience makes them experience anxiety without any other cognitive elements coming into play. The only reasons children abide by moral principles are instrumental; they act according to moral rules because they want to avoid unpleasant feelings and punishments.

As time passes, our personal reasons for acting in accordance with moral prescriptions (eg, our instrumental reasons) are replaced with objectified moral principles; principles of action that we reify from our proto-moral emotional experiences, but that we now understand as holding independently of those sentiments. Once these principles are established, moralizers no longer need to experience moral emotions in order to comprehend the legislative function of moral pronouncements. This characteristic of moral discourse is implicit within the way we employ morality. When moralizers use moral predicates, they assume that other moralizers will recognize that the use of these predicates suggest that the individual using them *demand*s assent regardless of the instrumental reasons others might have for doing so. In other words, we assume

---

<sup>33</sup> Raine et al, p. 210

that moralizers have more than instrumental reasons for acting in accordance with our moral prescriptions.

Those individuals who do not undergo proto-moral experiences fail to be able to comprehend what distinguishes moral reasons for acting from practical reasons for acting. Individuals with this type of deficiency in moral cognition are incapable of identifying the unique legislative function of moral assertions. For the psychopath, there are only instrumental reasons for acting morally; they perceive morality as most children perceive talking in class. The strongest reason (and in many cases, probably the only reason) children avoid talking in class is that they fear being punished; the fact that the rule has to be so rigidly enforced with threats of detention demonstrates that most children are not inclined towards avoiding disrupting class out of respect for their peers or other selfless reasons. These individuals act and reason as a metaethical instrumentalist might describe moralizers; the principles of action they accept are based on the practicality of committing themselves to such principles. A psychopath can act in accordance with moral principles, but the individual in question will never be capable of doing so for *genuinely moral reasons*. Recognition of the distinction between moral and practical reasons is one of the major capacities that distinguishes moralizers from psychopaths.

At this point in my investigation, it should be clear how the legislative function of moral discourse causes a problem for hermeneutical moral fictionalism. Insofar as a moral fictionalist might try to describe moralizers as individuals who act on moral principles only because of the instrumental value that follows from the acceptance of moral precommitments, this fictionalist fails to describe genuinely moral behaviour. The

behaviour she describes is that of a psychopath; one who cannot comprehend the unique magnetic-legislative function of moral discourse. This is not how moralizers understand their moral pronouncements.

When engaging in moral discourse, our hope is that other moralizers will be driven to act in accordance with our moral pronouncements whether or not they have instrumental reasons for doing so. This assumption on the part of moralizers is the antithesis of how metaethical instrumentalism would depict moral deliberation and moral commitment. It simply is not the case that moralizers assent to moral principles and precommitments because of the instrumentality of doing so. Moralizers obey moral principles because they understand them as being legislative; people are supposed to act in accordance with them regardless of how well they facilitate the long or short-term goals of those individuals involved.

It is clear, then, that the metaethical instrumentalist mischaracterizes the way we currently understand moral discourse. Only individuals with psychopathic symptoms actually understand moral commitments as they are depicted by hermeneutical moral fictionalism. But any criticism of hermeneutical moral fictionalism automatically raises questions for revolutionary fictionalist accounts. If a shift towards fictionalism would create a moral structure that cannot support some important aspects of our moral practice, the revolutionary fictionalist should address this fact and explain why the shift towards fictionalism is worth threatening the particular moral institution. In the following section of this paper, I argue that the inability of a fictionalist society to support the legislative function of moral discourse raises questions about the nature of moral imputation in a fictionalist moral community.

## VI: Moral Imputation and Revolutionary Fictionalism

The legislative function of moral practice facilitates stronger group cohesion and the ability on the part of moralizers to take action *against* those who disagree with or disregard their moral principles. This is an important element of moral practice; in fact, the attempt to account for this function is arguably one of the most important elements of normative moral theory and metaethical debate. Problems pertaining to ethical pluralism are widespread in the history of moral philosophy. Even today, there has been a significant amount of literature written on the topic of how the tenets of a Liberal society could be reconciled with radically alternative moral systems like those of Aboriginal Nations.<sup>34</sup>

In the absence of the magnetic-legislative character of correct moral claims, and lacking the ability to reconcile the similarities between subjective evaluative domains (such as preferences for food) and moral practice, we have no more reason to require that someone acts in accordance with our moral beliefs than we do to require that they enjoy the same types of food as we do. Theories that fail to vindicate this legislative function undermine moralizers' ability to legislate moral obligations to individuals who hold distinct, irreconcilable moral commitments. Lacking any means of preserving this magnetic-legislative functionality would entail the failure to differentiate between morality and other evaluative domains of discourse.

Accounts of moral practice have to be able to give reasons why moral pronouncements still apply to those individuals who have no personal commitment to those pronouncements. Even if an individual (Mrs. X) cannot foresee the instrumental

---

<sup>34</sup> See Rawls 1999a



value of a moral practice and has no personal commitment to said practice, she is still expected to act in accordance with it. The nature of moral discourse ensures that any moral proposition that would apply to another individual (Mrs. Y) who is in the exact same circumstances as Mrs. X would also apply to Mrs. X, regardless of their divergent personal commitments. Moral pronouncements are treated as if it is irrelevant whether or not any individual feels their magnetic pull at any particular time; if I believe that stealing is morally wrong, the nature of moral beliefs dictates that even if I ceased to experience any emotional response to the prospect of stealing, I would still consider stealing to be wrong. It may seem at first glance that this characterization of moralizers does not include utilitarians and other act-based consequentialist theorists. A utilitarian, for example, may not always agree that stealing is wrong. She may only conclude that *I* should not steal, because I have committed myself to the moral principle that rejects stealing; she would have no problem saying that other people should steal in certain circumstances. But even within a consequentialist society, there are moral principles that people are expected to obey regardless of their personal commitments. The best example of this from utilitarianism is the Happiness Principle; one must always aim to do what best facilitates the production of happiness. Regardless of any individual's personal commitments or desires, all moral persons are expected to uphold this principle. Moral principles like these are meant to apply to all individuals, regardless of motivational influences that might give some people reasons to violate them.

Metaethical instrumentalism requires that in our most critical metaethical contexts, we concede that our only reason for engaging in moral practice is that it best

helps the individuals in a community achieve their personal goals.<sup>35</sup> But when we couple this metaethical practice with the claims made in the earlier sections of this paper, most pertinently that we generally presuppose of any moral agent that she has the capacity to enter her most critical context and engage in metaethical deliberation, we run into certain problems. Predominantly, we come to find that all individuals will, at some point in their lives, be able to understand their moral practice as grounded in fundamentally egoistic concerns. The problem with this is that it is possible for individuals to accurately calculate that in some instances, disobeying the moral law will best serve their short-term *and* long-term interests; especially in cases wherein they can be guaranteed that none of their confederates will learn of their indiscretion.

This is the familiar problem of Hume's sensible knave and Hobbes' fool; the individuals wonder why they should cooperate unconditionally when they can receive the benefits of cooperation regardless of their own actions in a particular circumstance. But, in order to gain the full advantages of cooperation, every individual within the moral system would have to be assured that all *others* did not foresee the same dilemma as did the knave. This problem of the sensible knave is one that has been discussed at great length and in various capacities. My own invocation of this topic is motivated by my goal to demonstrate the inability on the part of metaethical instrumentalism to support our current practice concerning attributions of moral blame. Metaethical instrumentalism and moral fictionalism depict moralizers as little more than sensible knaves, acting for practical reasons instead of genuinely moral reasons. The problem with the presupposition that moralizers can be accurately characterized as sensible knaves is that it

---

<sup>35</sup> See Joyce (2001, 2005)

threatens our ability to hold individuals responsible to *our* moral commitments when these individuals have divergent commitments, goals and beliefs.

This is one of the major problems with the instrumentalist conception of morality. Since instrumentalism requires individuals to be motivated only out of personal interest, one only obeys moral principles because doing so facilitates the achievement of long-term goals. Recall that Joyce's depiction of moral fictionalists suggests that moral agents within a fictitious moral system may even *always* act consistently with the moral principles, despite their accurate calculation that violation of the principles in some particular instance would best serve the relevant interests. A moralizer may not always be able to perfectly calculate all the consequences of each of her actions; as such, at some point she may engage in a selfish, immoral act, believing that her confederates will not discover her, only to realize later that she had miscalculated. Consequently, she would be found out as violator of moral principles and would lose the long-term benefits of engagement within a moral society. Fictionalism can, then, still provide reasons to act morally whether or not doing so best facilitates short-term goals, but the reasons will still always have to do with acquiring instrumental benefits.<sup>36</sup>

If we were all to consent, in our most critical context, to such a purely instrumental explanation of our moral practice, (i.e., knowing on some level that most individuals will act morally only because it is in their best interests to do so) then when someone diverges from this norm our only means of criticizing their behaviour is via their inability to adequately calculate the course of action that best facilitates their personal advantage. But this is not what we do when we appeal to moral legislations;

---

<sup>36</sup> See Nolan et al (2002)

moral legislations are meant to be forceful independently of one's personal reasons for acting in accordance with them.

If a moral fictionalist (Y) were to conclude that an individual (X) ought not to hurt other people, she would have to be able to justify the moral system through which this normative claim could be established. Y's only explanatory recourse under an instrumentalist metaethic would be to attempt to explain how X's actions do not best preserve X's long-term goals; even risking being caught performing such an action would undermine X's own hypothetical imperatives. If, however, X were to have different motivational principles than Y, Y's moral claim should no longer be expected to have any magnetic effect on X.

Allow me to elaborate with an example. For most adult humans, going to prison is not a favourable option. Avoiding prison is one of the major instrumental benefits of obeying moral principles (at least, in a society with a reasonable match between moral and legal rules). But presuppose for a moment that there is some individual for whom going to prison *is* a favourable choice. Perhaps all of her friends and family are in prison for a long period of time, and she herself has no house, money or prospects as a free individual. She could get a low-income job and attempt to support herself, but because of the low wages, long hours and high cost of living, she finds this option extremely unfavourable. In prison, however, she would not only be surrounded by friends, but she would also be protected from other inmates by her allies, fed by taxpayers' money, sheltered and so on. As such, it is in her best interest to be identified by the society at large as a violator of moral principles; if she commits a violent crime and is recognized as

doing so then she will go to jail, which is in favour of both her long and short-term interests.

The instrumental metaethicist has no way to morally criticize this criminal for her action. If we acknowledge that only instrumental reasons underpin our moral commitments, we lose all capacity to morally impugn this person for her actions. Insofar as she has not acted against her best interests in either the long or short-term, how could she have benefited from engaging in moral behaviour? If our only reasons for not engaging in immoral behaviour are because of our desire to be trusted by our confederates, then the normative system upheld by our instrumental metaethic has no justifiable hold on her. In being found out as an individual that takes advantage of a cooperative society, she actually betters her circumstances. She would also have no use for Joyce's instrumental moral heuristic, since being discovered as a violator of moral principles best serves her interests.

The argument might be raised that it is only in our most critical context that the moral fictionalist must admit there is not anything wrong with the criminal's action; in all other contexts she could presumably continue to employ her functional moral discourse. When we examine the actions of any individual in our most critical context, instrumental metaethical theories would have us believe that *no* actions could properly be called right or wrong; in critical contexts we vitiate all the moral assertions that we assent to in other contexts. Thus, for me to say that we lack all ability to morally impugn this individual in our most critical context has no force because, according to instrumental metaethical theories, we cannot morally impugn *anyone* from this context.

But this argument misses the point. Any individual could be said to have *reasons* to act morally, even if her actions were to be examined in the most critical context. If these individuals wanted to gain from the benefits of social interaction, they ought to conform to moral norms. If the antecedent of this conditional fails to obtain, then it is clear that the individual in question has no reason to be moral. If an individual *does* want to gain the benefits of social interaction and yet still commits a violent crime, the instrumental metaethicist still has reason to say that this individual has done something *wrong*. For the instrumental metaethicist, there is still a relevant difference between the two individuals; one should have engaged in the violent act while the other should not have.

This account paints a picture of moral judgement that is not only highly inaccurate, but also quite troubling. Presumably, any individual would be capable of slipping in and out of her most critical context at virtually any time if she was pressed on the right questions. If this person were an instrumental metaethicist, then she would impugn the violent criminal described above as would any other moralizer. If I were, however, to then drag this person into a philosophy classroom and ask her what she would have done in the criminal's position, she would have to respond in terms relative to the criminal's instrumental reasons for acting (since she would be in a philosophy classroom at the time I posed the question, her answer would, *prima facie*, have to be a product of metaethical deliberation in her most critical context). Instrumental metaethics thereby paints a picture of a moralizer as someone who asserts that an individual is morally deplorable, yet concedes that barring certain contingent circumstances of her own life, she would have and, this is the truly troubling part, *should have* acted in an

identical fashion. The obligation in this claim is one of pragmatic and practical normativity, not moral normativity, but this point still raises questions about instrumental metaethical theories. These theories suggest that moral considerations entail that the criminal should have refrained from acting, yet metaethical considerations entail that her action was not only permissible, but obligatory. This is part of the reason why the failure to preserve the magnetic-legislative function of moral discourse is problematic for the revolutionary moral fictionalist. Without offering a theoretical means via which we can preserve the distinction between moral and non-moral prescriptions and recommendations, the moral fictionalist recommends that a community of moralizers begins to openly acknowledge itself as a society of hypocrites; individuals who say one thing for the sake of pragmatic utility, but actually go about doing the exact opposite of what they recommend when their personal goals and motivations lead them in such a direction.

Instrumental metaethics, as it is employed by Joyce and other fictionalists, allows for moral behaviours to be justified and explained only in terms of the personal emotions and the competing desires that motivate any particular action. It thereby diminishes the distinction between genuinely moral assertions and beliefs (those being beliefs and assertions that lend themselves to magnetic-legislation) and personal evaluative assertions (such as a preference for a particular kind of food). Lacking this distinction, we are left with an extremely unfavourable depiction of moralizers as hypocrites; individuals having no scruples about condemning an action as deplorable and yet validating the rational deliberation that motivated the act. This depiction of moralizers undermines our concept of moral imputation, since instrumentalist theories require that the moral value of

any individual's action is relative to that individual's set of goals and desires; no act performed by an individual is wrong as long as it was the best thing she could have done relative to her own interests.

If accepting metaethical instrumentalism leaves us with no option but to concede that different people's actions should be held to different standards of judgement, even if only in our most critical contexts, then this practice threatens the magnetic legislative function of moral discourse. The consequences of the inability to coherently support this function of moral practice are immense. Universality and impartiality, common themes in nearly all normative theories, could only be acknowledged as pragmatic tools that have no real bearing on what one ought to do. Virtually every influential normative theory recognizes and attempts to explain why universality and impartiality are relevant facets of moral practice. Utilitarianism emphasizes the importance of maximizing global utility, not just personal utility, and Kantian deontology considers a lack of concern regarding one's own circumstances essential to any truly moral action. If revolutionary moral fictionalism fails to preserve the universal impartiality of moral discourse, it fails to capture one of the most important distinguishing aspects of morality.

Another aspect of moral discourse that would be lost in a fictionalist revision is the supervenience of moral predicates. Supervenience is the property of moral predicates that ensures that the instantiation of a predicate in one situation requires the instantiation of the same predicate in all relevantly similar situations. Blackburn defines supervenience in the following way: "A Property  $M$  is supervenient upon properties  $N_1 \dots N_n$  if  $M$  is not identical with any of  $N_1 \dots N_n$  nor with any truth function of them, and it is logically impossible that a thing should become  $M$  or cease to be  $M$  or become more or less  $M$  than



before, without changing in respect of some member of  $N1 \dots Nn$ .”<sup>37</sup> Moral properties, like goodness, rightness and wrongness supervene upon otherwise value-neutral characteristics of states of affairs. If two circumstances share all the same non-moral properties that are relevant to the attribution of a moral predicate, they should share all moral properties.

This function of moral discourse is instrumentally valuable for a moral community. Without the supervenience of moral predicates, we lose all capacity to morally criticize individuals for engaging in acts that other moralizers would have found reprehensible. Moral pronouncements could be rejected, not because of a change in morally-relevant properties, but because of other facts that were irrelevant to the original attribution of the moral predicate. Circumstances that were unrelated to a moral property could thereby influence people recognizing or denying instantiations of the property. The citizens of a moral society that did not employ supervenient moral predicates could therefore justifiably cite whatever reason they want, even short-term egoistic goals, as relevant reasons for neglecting a moral duty. Such a society would, in other words, allow for conflicting moral predicates to be applied to one set of morally-relevant properties, based on properties that *should* be excluded from moral deliberation.

Since metaethical instrumentalism requires that there be different standards of conduct applying to people with divergent sets of motivational influences, moral properties in a fictionalist society would not be supervenient. Two individuals could be in situations that are identical in terms of all relevant moral properties but, because of each of their own instrumental reasons (granting that one’s own personal non-moral desires are

---

<sup>37</sup> See Blackburn 1993(c) for a detailed account of how supervenience is a relevant feature of moral practice.

not morally relevant), have divergent obligations. A moral fictionalist society thereby fails to preserve virtually any of the legislative functions of moral discourse.

Without our ability to hold others to impartial, universal, supervenient moral principles, we lose many of the instrumentally valuable functions of morality. The revolutionary moral fictionalist is quick to demonstrate those aspects of morality that could potentially be preserved by a society-wide shift to fictionalism, without weighing the possible costs in the balance. Presumably, if the erroneous foundations of morality could be accounted for without having to sacrifice the legislative function of moral discourse and our normative concept of moral personhood, the fictionalist would be in favour of this option over her own proposal. The final section of this essay is concerned with the possibility that the fictionalist herself would prefer the success of a prescriptive quasi-realist project over the success of her own. In this way, her metaethical strategy is mistaken because her own criterion recommends another candidate.

## VII: The Economy of Metaethical Instrumentalism: Quasi-Realism over Moral Fictionalism

When we face the doubt that arises as a response to some of our more questionable realistic moral practices, we actually have a tangible option between at least two distinct approaches; moral fictionalism as it was presented thus far, and another project known as *quasi-realism*. The remainder of this paper is devoted to arguing that a prescriptive quasi-realism is a more attractive alternative to moral fictionalism, in that it could be employed to respond to the worries of Joyce and other error-theorists without threatening some of the most valuable aspects of our moral practice. A moral fictionalist, in order to make a compelling case that fictionalism is an attractive option, would first have to demonstrate or at least strongly suggest the failure of the quasi-realist project.

The form of quasi-realism that will be addressed in this essay is largely extrapolated from the work of Simon Blackburn. There are other forms of quasi-realism that stand largely independently of Blackburn's specific considerations<sup>38</sup>, but explanation of Blackburn's general position and the prescriptive project that could be extrapolated from his work will suffice for the task at hand.

Joyce does offer a brief analysis of how quasi-realism diverges from the goals of his own project.<sup>39</sup> His only argument against quasi-realism is that the fictionalist presupposes its failure due to her previous acceptance of a moral error theory. The type of quasi-realism I am recommending would, however, be a *revolutionary* quasi-realism, that is, an understanding of moral language that we *could* assent to upon learning that realist

---

<sup>38</sup> See predominantly Gibbard (1990, 2003)

<sup>39</sup> See Joyce (2001), p. 201

moral discourse must be abandoned. Since this type of quasi-realism is addressing the same possibility as is the fictionalist, Joyce has no justifiable reason to presuppose that a shift towards quasi-realist discourse could not be successful. Thus, fictionalists should treat quasi-realism as a live option.

Understanding the quasi-realist project requires the explanation of three distinct yet related philosophical positions; expressivism, projectivism and quasi-realism. Expressivism is the heir to classical non-cognitivism about moral language. Rather than focusing on the non-assertoric nature of moral utterances or the relationship between such utterances and prescriptions for action, contemporary expressivism is best defined as an explanatory stance towards moral or evaluative beliefs. Under the contemporary expressivist reading, for S to believe that X is good is just for S to value X.<sup>40</sup> An assertion of the goodness of X can thereby be explained in terms of S' attitude or stance towards X. This not to say, however, that S' assertion of the goodness of X is S *describing* her own state of mind or stance towards X, but rather that she is *expressing* it.

Projectivism is an explanatory framework that Blackburn adapts from David Hume. A projection occurs when an attitude, habit or other commitment which is evaluative, expressive or otherwise non-descriptive is spoken, thought and reasoned about as if it were descriptive. We presuppose, with projective discourses, that the objects of the discourses can be properly treated *as if* they were descriptive in the way that we employ them. Blackburn and others often characterize projective properties as the mind *spreading itself* on the world.<sup>41</sup> In short, our moral reactions, attitudes and expressions come to be projected upon the world and thus the sentiments that underpin these

---

<sup>40</sup> See Blackburn 1998, p. 50

<sup>41</sup> See Blackburn (1984), p. 171

projections lead to beliefs about the nature of moral predicates that resemble those of a moral realist. Projectivism could thus be seen as a causal explanation of how we come to have the type of moral beliefs that we do.

The project of quasi-realism centres on the acceptance of the expressive nature of moral utterances in conjunction with the attempt to establish that we are not mistaken about the way projective predicates should be used. Blackburn characterizes the quasi-realist project as the attempt to earn the right to use realistic moral discourse despite acceptance of an anti-realist metaethical starting point.<sup>42</sup>

Prescriptive, or revolutionary quasi-realism (that being the suggestion that we come to employ morality in such a way that the quasi-realist's project would be successful) would bear several similarities to revolutionary moral fictionalism. Both would acknowledge an anti-realist metaphysics regarding moral facts and beliefs and both would attempt to offer a vindication of some aspects of our realistic moral practice. David Lewis actually goes so far as to say that quasi-realism is a variety of moral fictionalism; insofar as the 'realistic' moral assertions of the quasi-realist can actually be characterized as little more than 'quasi-assertions', everything a quasi-realist says about morality is supposed to be supplemented by some assumptions about the anti-realistic nature of moral discourse. The real content of these assertions is dependent upon the non-realist caveat that projectivism makes us speak 'as if' moral predicates can and should be used descriptively and realistically.<sup>43</sup>

Blackburn is correct to differentiate his own philosophical position from that of the moral fictionalist. Quasi-realism, unlike moral fictionalism, seeks to account for the

---

<sup>42</sup> Ibid

<sup>43</sup> See Lewis, p. 319

anti-realist problems *without* invoking metaethical instrumentalism as an argumentative strategy. Rather than characterizing moralizers as individuals who purposefully adopt moral precommitments because of the instrumental value of doing so, quasi-realism retains the realist approach of attempting to demonstrate that moral precommitments, beliefs and facts are functionally equivalent to facts and beliefs that are *literally* true. This literal truth comes in the form of quasi-truth, which will be discussed more in the paragraphs to come, but the distinction between fictionalism and quasi-realism is readily apparent, in that the quasi-realist allows for us to engage in moral discourse because moral discourse is, in some sense, *true*, rather than just instrumentally valuable.

There is also good reason to suspect that the nature of the quasi-realist project actually lends itself to the preservation of realistic practice much more effectively and economically than does the moral fictionalist. If we were to affirm a prescriptive quasi-realist theory, that being one that suggests we fully enumerate and classify the nature of moral projections and thus earn the right to discuss projected predicates as if they were *not* factually defective, we would not run afoul of the unfavourable characteristics of a revolutionary moral fictionalism that was depicted in earlier chapters.

Consider, for a moment, the conceptual moves made by the moral fictionalist. Upon realizing that there is something wrong with the way we employ moral discourse (e.g., that we use the terms realistically when the instantiation of moral predicates actually relies on attitudes and stances) and accepting that this problem is a fatal one for any justificatory account of moral discourse, we have no choice but to either abandon morality or adopt a fictionalist attitude towards it. Seeing that we preserve some benefits with fictionalism, it is better to accept this account than to favour abolition. The move to

fictionalism requires recognition that in order to preserve the functionality of moral discourse, we have to change the way we use it; we make what Joyce considers to be a move towards non-cognitivism.<sup>44</sup> We begin to regard moral practice as a matter of employing useful heuristics, thoughts and other cognitive activities in order to facilitate cooperation and avoid akrasia. The moral fictionalist recognizes that any of her cognitivist commitments were actually a product of erroneous assumptions on her part.

It is important to note that the prescriptive quasi-realist and the revolutionary moral fictionalist assent to a number of similar claims. For example, a prescriptive quasi-realism would claim that we should make a move towards non-cognitivism and that this move towards non-cognitivism facilitates employing moral discourse *as if* there was nothing wrong with our realistic practices. It is this kind of articulation of the project that leads some theorists to relate these two distinct philosophical projects to one another. But there are actually significant differences between fictionalism and quasi-realism. As Blackburn himself points out, there is a relevant distinction between how the two positions employ the phrase ‘as if’.

The fictionalist speaks ‘as if’ there are real moral properties while theoretically assenting to the fact that such properties only exist in terms of the fiction. But this is not the only way we can be said to speak ‘as if’ realistic moral practices are justifiable. We can also be said to be speaking ‘as if’ we are expressing real attitudes in a language that acknowledges and is well-suited to coping with this fact.<sup>45</sup> But since this is what we actually do when we engage in moral discourse, why should we be characterized as pretending or engaging in ‘make believe’ at all? The fictionalist asks us to pretend that

---

<sup>44</sup> See Joyce (2001), p. 200

<sup>45</sup> See Blackburn (2005), p. 329

moral discourse is factual in order to gain some kind of benefit, but the quasi-realist argues that when we treat moral discourse factually, we are actually employing a practice that was earned via an enumeration and analysis of projected properties. Quasi-realism does not make us pretend, it demonstrates that we *don't have to* pretend.

Another way to characterize the distinction between moral fictionalism and quasi-realism goes hand in hand with my attribution of the burden of proof onto the moral fictionalists. If we were able to show, via an enumeration of cross-cultural projective tendencies, that quasi-realism could be a successful project (which I suspect it might be), we would not be relying on solely instrumental reasons for engaging in moral behaviour. Since quasi-realistic moral discourse earns all aspects of realistic moral discourse, quasi-realists have no need to appeal to the instrumental value of moral practice in order to vindicate engagement in said practice. There is no longer any need to justify our moral precommitments on the basis of their instrumental value because quasi-realism shows us that we were not *really* in error at all when we presupposed the existence of realistic moral facts; we just thought we were in error because we misunderstood what we were doing. The fact that quasi-realism can preserve our realistic reasons for engaging in moral discourse, (eg, that the truth of moral claims is not grounded in the instrumental value of assenting to said claims) serves to further differentiate the quasi-realist from the fictionalist. It also suggests that quasi-realism would be able to preserve more instrumental benefits of moral discourse than could fictionalism.

Part of what a shift to quasi-realism would constitute is the recognition that all of our realistic moral tendencies are in fact entirely justifiable for *more* than just instrumental reasons. The discourse would thereby be 'sanitized' from the supposition



that certain errors are latent within any application of moral predicates. Insofar as our common realistic practices are liberated from the unfavourable diagnosis offered by most anti-realists, all realist reasons we may have had for engaging in moral practice are thus preserved. If, as I have thus far contended, moral realists have reasons for engaging in moral discourse that go beyond the instrumental uses for moral practice, then the quasi-realist would allow us to continue to cite these reasons when in our most critical metaethical contexts.

Quasi-realism, then, is no more committed to the instrumentality of moral discourse than is realism. The functions of moral discourse that I outlined earlier in this essay would thereby be preserved by a move towards quasi-realism, while being threatened by a move towards fictionalism. Whereas revolutionary moral fictionalism threatens these functions in the attempt to preserve certain other useful mental practices and tendencies, revolutionary quasi-realism preserves everything about moral discourse.

So why might the moral fictionalist have to justify the adoption of a fictionalist stance over a quasi-realist one? The moral fictionalist has only one criterion with which to decide what type of moral practice to adopt. This criterion relies entirely on the instrumentality of that practice. She only endorses fictionalism because of the instrumental values that could be preserved in a fictionalist society. If I am right that quasi-realism manages to preserve *more* instrumental benefits than does moral fictionalism, then the moral fictionalist should herself prefer a quasi-realist theory *unless* she can demonstrate some errors that are implicit in an acceptance of quasi-realism. This is simply because the instrumentality of quasi-realism exceeds that of fictionalism; it preserves more functions of moral discourse that have direct instrumental benefits. Since

the quasi-realist allows for assertability conditions of projective moral predicates to confer 'quasi-truth' on moral claims, quasi-realism offers no challenge to the magnetic-legislative function of moral discourse.<sup>46</sup> Insofar as the moral predicates are projected onto the world, the quasi-realist would hold not only that they stand apart from other evaluations or desire-relative hypothetical imperatives, but also that their unique legislative properties are justifiably a part of the assertoric content of moral utterances. Preserving this function allows for the quasi-realist to continue to employ legislative moral practices without relying on instrumental reasons for doing so.

Additionally, quasi-realism facilitates the understanding of metaethical discourse as being strongly interrelated with our first-order moral deliberations; this lends itself to a favourable depiction of moral persons as those individuals capable of reconciling metaethical and moral beliefs. This point may be somewhat controversial, as Blackburn has rallied against claims that only 'valuers can understand values' or that there is no means by which to offer an external explanation of morality.<sup>47</sup> But quasi-realism is in fact more compatible with this facet of moral discourse than it may superficially appear.

Quasi-realism takes a fairly unique perspective towards the truth of the thesis of moral projectivism. Many realists treat projective theories as taking a reductive stance towards moral discourse or rendering all morality relativistic.<sup>48</sup> But the attitude quasi-realists would take towards projectivism is one that allows for the instantiation of projective predicates to have assertability conditions and thus quasi-truth conditions. This

---

<sup>46</sup> Blackburn, in many occasions, makes it known that he considers this aspect of moral discourse to be fundamental to distinguishing between moral assertions and other evaluations; "...the strongest ethical judgements do not issue from stances that are properly variable" (Blackburn, 1993(a) p. 178). Moral evaluations, for the quasi-realist, are distinct from all other types of evaluations.

<sup>47</sup> See Blackburn 1993(a), p. 177

<sup>48</sup> Ibid

allows for substantial moral debates to occur between rational individuals; quasi-realist conceptions of projection do not threaten the substantive content of moral claims nor do they deflate the existence of actual moral controversy. Everything that realistic moral practice assumes about moralizers is earned by the quasi-realist.

Recall that one of the primary problems with instrumentalist metaethical theories is that they threaten our conception of rational moralizers as potentially ‘ideally critical’ metaethicists; they do not allow for the fact that all moral persons must be able to offer non-instrumental reasons for engaging in moral practice. Whereas reliance on only instrumental reasons threatens moralizers’ ability to adjudicate between good and bad reasons for adopting any particular moral practice, realistic metaethical deliberation does not. The fictionalists’ only recourse when faced with the problems realistic discourse has accounting for the justification of this practice is to offer some instrumental reason to preserve the particular fiction of moral personhood. But, as I argued earlier, no answer that the fictionalist can offer to the question of why we should engage in morality sufficiently preserves the concept of a moral person, which is not only central to our ‘out-group’ mode of moralizing, but also helps us identify those individuals who are capable of genuine moral behaviour.

Quasi-realism, however, preserves this type of rigorous metaethical debate. A quasi-realist, in her most critical context, would be able to offer normative justifications for preferring some reasons to others (e.g., she would be able to say that her society should not adopt a Spartan ‘in-group’ moral system, even though this system yields the greatest instrumental value for all involved) and would be able to preserve the realist’s normative concept of a moral person, which is such an important aspect of our current

moral practice. For quasi-realists like Blackburn, moralizing is comparable to Neurath's boat; metaethical considerations cause revisions of some other aspects of the practice, but the practice is not relativistic because any moral or metaethical question (i.e., any assertion employing the predicates that stem from projections) is necessarily performed on the boat. Once you step off the boat, you cease to engage in ethics. This is a more favourable depiction of moral agents that matches up more effectively with some important features of our moral practice.

Before the moral fictionalist can make a revolutionary claim like Joyce's, she would thereby first have to provide reasons why quasi-realism is not a viable project. Even though quasi-realism is superficially similar to fictionalism, the two projects stand apart insofar as the conceptual shift prescribed by quasi-realism would not have us significantly alter the function of moral discourse. Prescriptive quasi-realism would seek to preserve all functions of the discourse despite the acceptance of anti-realist presumptions. Since it does not rely on only instrumental reasons for preserving moral discourse and practice, quasi-realism does not fall prey to the same dysfunctions as would a fictionalist moral system. But this fact in turn actually renders a quasi-realist moral system *more* instrumentally useful than would be a fictionalist system; it preserves all the same advantages without threatening the important normative institutions on which we moralizers rely. If the moral fictionalist is truly concerned with the instrumentality of moral practice, she should hope for the viability of prescriptive quasi-realism and only rely on moral fictionalism pending the failure of said project.

There may in fact be some reasons, however, why a quasi-realist system may be less favourable than a fictionalist one. One such reason is the problem of propositional

attitudes; non-cognitivists or expressivists have difficulty justifying how, exactly, we might go about accounting for the way we talk and reason about the stances people take towards moral properties. One of the most central problems is the issue of belief, but the same problem arises any other time we try to discuss a moral attitude without expressing or asserting it.

If all moral assertions are actually non-cognitive or expressive, then any time I express a moral commitment, I express to others the stance I take towards the object of my assertion. As such, non-cognitivists have difficulty coping with statements such as ‘Mary believes that abortion is wrong’.<sup>49</sup> If our moral assertions actually are just non-cognitive expressions of our attitudes towards objects, how are we to characterize a moral attitude without expressing it ourselves? Since the meaning of ‘abortion is wrong’ remains the same whether I express my own attitude or describe someone else’s, some theorists have suggested that expressivism does not adequately explain the semantic content of moral utterances.

These types of problems have been addressed as embedding problems, or have sometimes been referred to as ‘Frege-Geach’ style problems, after the two theorists who addressed the problem most famously. The problem relates predominantly to how expressivism could cope with unasserted uses of an expressive phrase. One example comes in the form of arguments involving moral predicates like the following:

- P1: If stealing is wrong, then Johnny should not steal
- P2: Stealing is wrong
- C: Johnny should not steal

The problem with these unasserted contexts is that expressivism holds that moral utterances express attitudes and do not represent propositions. Since in the first premise,

---

<sup>49</sup> Nolan et al, p. 25

the phrase 'stealing is wrong' is not asserted but is rather embedded within a conditional, this phrase does not express the attitude towards stealing that is present in premise two. The meaning of 'stealing is wrong' thereby diverges in each premise. But, as the opponents of expressivism point out, the meaning *is* the same in each of the premises. The same is true for the case of propositional attitudes like belief, which I addressed above. If expressivism denies that moral claims represent propositions, and propositions are the objects of belief, how can we account for embedded moral phrases? This is one of the most difficult problems for proponents of expressivism. Moral fictionalists could, however, claim that their conception of moral propositional attitudes fares better than does that of the quasi-realist because fictionalists can continue to presuppose that moral assertions are genuinely *assertoric*, but adulterated by the fact the content of the assertion is grounded in a fictive discourse.

Quasi-realists like Blackburn have attempted to explain away this particular problem,<sup>50</sup> but there are also reasons to question the moral fictionalists' claim that they actually do preserve the assertoric content of moral utterances. My reasons for suggesting this stems predominantly from some of the considerations Joyce includes in his account of revolutionary fictionalism about morality. He points out that for a fictionalist moral system to develop, we cannot think about our moral commitments in terms of beliefs.<sup>51</sup>

Strictly speaking, moral fictionalists do not *believe* the moral propositions they employ. In their most critical context, they vitiate all of their previous moral assertions and professed commitments. One problem that arises out of this failure to substantiate moral beliefs is that it weakens the moral commitments individuals could discuss. On

---

<sup>50</sup> See Blackburn (1984), pp. 189-196

<sup>51</sup> Joyce (2001) pp. 200-201

certain occasions, individuals might *act* as if they believe in certain obligations and rules, but this would only be because they are using the purported belief as a heuristic to facilitate the acquisition of some other goal.

Joyce characterizes this fictionalist moral practice as the employment of *thoughts*, not beliefs.<sup>52</sup> Referring back to the example of performing a certain number of sit-ups, one might think to oneself ‘must...do...fifty’, when in fact one could admit to oneself at another time that some days forty would be as good as fifty. It is the *thought* that is instrumentally useful at a particular time, not the belief.

Aside from the obvious complications regarding when we decide to invoke the thought (e.g., why could I not, while doing my fortieth sit-up, conclude that forty is actually sufficient?), this also causes a major problem for Joyce’s fictionalism. One of the major benefits some other authors have attributed to fictionalism is its capacity to account for the quasi-realists’ problems regarding propositional attitudes.<sup>53</sup> But if we must do away with the discourse of moral belief, the fictionalist does no better than does the quasi-realist. Since Joyce endorses an attitude towards moral language that relies on a move towards non-cognitivism, he would have to explain away the same type of Frege-Geach style problems that the prescriptive quasi-realist would be forced to face. For example, when fictionalists utter the phrase ‘stealing is wrong’, they express their precommitment to the fictive moral principle. When they utter ‘if stealing is wrong, then Johnny should not steal’ they do not express a precommitment, suggesting that the meaning of the phrase changes in each premise.

---

<sup>52</sup>Ibid

<sup>53</sup>Nolan et al (2002), p. 25

The ‘thoughts’ that Joyce describes might, however, be somewhat reconcilable with the problem of propositional attitudes. One might argue that these thoughts, insofar as they are fictional claims, are propositional, but factually defective.<sup>54</sup> The truth ascribed to these propositions is adulterated by the fact that it relates to fictional facts, but the moral fictionalists’ discourse could presumably remain genuinely representative of propositions. Thus, the adulterated truth of any fictional obligation could still be preserved in logical inferences or attributions of belief. This may be true for fictionalism in general; in fact, this is the exact route taken by Nolan in his own defence of moral fictionalism against quasi-realism.<sup>55</sup> My argument, however, is geared specifically towards Joyce’s articulation of how revolutionary moral fictionalism might best be applied. Insofar as Joyce recommends that a shift towards fictionalism requires the adoption of a form of expressivism, his revolutionary fictionalism does not evade this particular problem.

Whether or not quasi-realism is a viable project is an important question, the scope of which extends well beyond the subject matter of this essay. All I hope to demonstrate, in comparing the virtues and vices of moral fictionalism and quasi-realism is that the moral fictionalist’s own metaethical standards imply that revolutionary quasi-realism would be preferable over a conceptual shift towards a Joycean society of moral fictionalists. Unless Joyce were to offer some reason why the quasi-realist project should come under suspect, he could not continue to claim that moral fictionalism is the most preferable option for moralizers forced to come to terms with moral anti-realism.

---

<sup>54</sup> This argument stems from comments received from Dave DeVidi.

<sup>55</sup> See Nolan et al (2002)



## **Conclusion**

Any metaethical theory that provides only instrumental reasons for why moralizers should continue to engage in moral discourse threatens the use of some normative concepts that are not only extremely useful, but important for the preservation of the peculiar functions of moral discourse. Revolutionary moral fictionalism, as articulated by Richard Joyce, succumbs to all of these criticisms. Thus, any instrumental benefit that can be acquired through a shift to moral fictionalism comes at the cost of some instrumentally valuable features of moral practice. Quasi-realism could serve as a more favourable, expedient alternative to revolutionary moral fictionalism and, as such, the moral fictionalists' own argumentative strategies require that they find fault with the quasi-realist project before she can persuasively endorse her own theoretical position. Failing to offer reasons why the success of quasi-realism should be doubted, there is no justifiable basis from which revolutionary moral fictionalism can be coherently recommended.

## References

- Blackburn, Simon. 'Quasi-Realism no Fictionalism'. From: *Fictionalism in Metaphysics*. Ed. Mark Eli Kalderon. Clarendon Press, Oxford. 2005. pp. 322-338
- Blackburn, Simon. *Ruling Passions*. Clarendon Press, Oxford. 1998.
- Blackburn, Simon. 'How to be an Ethical Anti-Realist'. From: *Essays in Quasi-realism*. Oxford University Press, Oxford. 1993(a). pp.
- Blackburn, Simon. 'Moral Realism'. From: *Essays in Quasi-realism*. Oxford University Press, Oxford. 1993(b). pp. 111-129
- Blackburn, Simon. 'Supervenience Revisited'. From: *Essays in Quasi-realism*. Oxford University Press, Oxford. 1993(c). pp. 130-148.
- Blackburn, Simon. *Spreading the Word*. Clarendon Press, Oxford. 1984.
- Blair, R.J.R. 'A Cognitive Developmental Approach to Morality: Investigating the Psychopath'. From: *Cognition* Vol. 57. Elsevier Science. 1995. pp. 1-29.
- Blair, James; Mitchell, Derek; Blair, Karina. *The Psychopath: Emotion and the Brain*. Blackwell Publishing, Oxford. 2005.
- Boyd, Richard. 'How to be a Moral Realist'. From: *Essays on Moral Realism*. Ed: Geoffrey Sayre-McCord. Cornell University Press, New York. 1988. pp. 181-118
- Burgess, John P. 'Why I am not a Nominalist'. *Notre Dame Journal of Formal Logic*. Vol. 24. 1983. pp. 93-105
- Dennett, Daniel. *Brainstorms*. MIT Press. Cambridge. 1981.
- Doris, John. 'Persons, Situations and Virtue Ethics'. *Noûs*, Vol. 32, No. 4, (Dec., 1998), pp. 504-530
- Dworkin, Ronald. 'Objectivity and Truth: You'd Better Believe It.' From: *Philosophy and Public Affairs*, Vol. 25, 1996. pp. 87-139.
- Field, Hartry. 'Disquotational Truth and Factually Defective Discourse'. *The Philosophical Review*. Vol. 103. No. 3. 1994. pp. 405-452
- Fields, Lloyd. 'Psychopathy, Other-Regarding Moral Beliefs, and Responsibility.' From: *Philosophy, Psychiatry & Psychology* Vol. 3.4. John Hopkins University Press. 1996. pp. 261-277

Frankfurt, Harry. 'Freedom of the Will and the Concept of a Person.' *The Journal of Philosophy*, Vol. 68, No. 1. 1971. pp. 5-20.

Garner, Richard & Rosen, Bernard. *Moral Philosophy: A Systematic Introduction to Normative Ethics and Meta-ethics*. The Macmillan Company, New York. 1967.

Gibbard, Allan. *Wise Choices, Apt Feelings*. Harvard University Press, Cambridge. 1990.

Gibbard, Allan. *Thinking How to Live*. Harvard University Press, Cambridge. 2003.

Hare, Robert D. *Without Conscience: The Disturbing World of the Psychopaths Among Us*. The Guilford Press, New York. 1999.

Hobbes, Thomas. *Leviathan*. Ed. Edwin Curley. Hackett Publishing Company, Inc. Cambridge. 1994.

Hume, David. *An Inquiry Concerning Human Understanding*. Library of Liberal Arts. New Jersey. 1995.

Joyce, Richard. 'Moral Fictionalism'. From: *Fictionalism in Metaphysics*. Ed. Mark Eli Kalderon. Claredon Press, Oxford. 2005. pp. 287-313

Joyce, Richard. *The Myth of Morality*. Cambridge University Press. Cambridge. 2001.

Kalderon, Mark Eli. *Moral Fictionalism*. Claredon Press. Oxford. 2005.

Kant, Immanuel. *Groundwork for the Metaphysics of Morals*. Tans: James W. Ellington. Hackett Publishing Company. Indianapolis. 1981.

Lewis, David. 'Quasi-Realism is Fictionalism'. From: *Fictionalism in Metaphysics*. Ed. Mark Eli Kalderon. Claredon Press, Oxford. 2005. pp. 314-321

Mackie, John. 'A Refutation of Morals.' *Australasian Journal of Philosophy*. 1977

Marino, Patricia. "Expressivism, Deflationism, and Correspondence," *The Journal of Moral Philosophy*, 2.2. 2005. 171-191.

McCloskey, H.J. *Meta-ethics and Normative Ethics*. Martinus Hijhoff, Netherlands. 1969.

McDowell, John. 'Values and Secondary Qualities'. From: *Essays on Moral Realism*. Ed: Geoffrey Sayre-McCord. Cornell University Press, New York. 1988. pp. 166-180.

Mill, John Stuart. *Utilitarianism*. Hackett Publishing Company, Indianapolis. 2001.

Nolan, Daniel; Restall, Greg; West, Caroline. 'Moral Fictionalism'. Draft only, cited with permission. (2002). The University of Melbourne Digital Repository. Available URL:  
[http://dtl.unimelb.edu.au/R/YIC1VLEL6J5Y8F6CJEKC36Q8TK9ED892XCAD7GC5XJTDN1MKPP-01459?func=dbin-jump-full&object\\_id=65833&pds\\_handle=GUEST](http://dtl.unimelb.edu.au/R/YIC1VLEL6J5Y8F6CJEKC36Q8TK9ED892XCAD7GC5XJTDN1MKPP-01459?func=dbin-jump-full&object_id=65833&pds_handle=GUEST). Date Accessed: July 5, 2008.

Rain, Adrian; Yang, Yaling. *Neural Foundations to Moral Reasoning and Antisocial Behaviour*. Oxford University Press, 2006.

Rawls, John. *A Theory of Justice*. Harvard University Press. Cambridge. 1999(a).

Rawls, John. *The Law of Peoples*. Harvard University Press. Cambridge. 1999(b).

Sayre-McCord, Geoffrey. 'The Many Moral Realisms'. From: *Essays on Moral Realism*. Ed: Geoffrey Sayre-McCord. Cornell University Press, New York. 1988. pp. 1-23

Stevenson, Charles Leslie. 'The Emotive Meaning of Ethical Terms'. From: *Mind*, Vol. 46, No. 181. 1937. pp. 14-31

Wiggins, David. 'Truth, Invention and the Meaning of Life'. From: *Essays on Moral Realism*. Ed: Geoffrey Sayre-McCord. Cornell University Press, New York. 1988. pp. 127-165.

Williams, Bernard. *Truth and Truthfulness: An Essay in Genealogy*. Princeton University Press. Princeton, NJ. 2002.

Williams, Bernard. 'Truth in Ethics'. From: *Truth in Ethics*. Ed. Brad Hooker. Blackwell Publishers, Oxford. 1996. pp. 19-34.

Williams, Bernard. 'A Critique of Utilitarianism'. From: *Utilitarianism, For & Against*. Cambridge University Press. Cambridge. 1973.

Yablo, S., 2001, 'Go Figure: A Path Through Fictionalism', in *Figurative Language: Midwest Studies in Philosophy*. Vol. 25. 2001, pp. 72-102.