How Emotivism Survives Immoralists, Irrationality, and Depression

Gunnar Björnsson
Stockholm University

When someone thinks that it is morally wrong to cut social welfare while increasing military spending, that state of mind is, essentially, something akin to a desire for such changes not to take place. Or so versions of the metaethical theory known as emotivism have it, claiming that moral opinions are best understood as motivational states.

Emotivism is hardly the received view of contemporary analytic philosophy, and much writing in metaethics seems bent on showing why and how this view is wrong and on producing something that has at least a hope of being right, that is, a theory denying that moral opinions are well understood as desires or motivational states. Perhaps the most popular objection to emotivism is the following:

*Immoralism*: If moral opinions are desires or similar motivational states, it is a mystery that there are what I will somewhat bluntly call ‘immoralists’: people having moral opinions but lacking the corresponding desires or motivation. (Some would prefer to label such people ‘amoralists’, but that label suggests someone who thinks that there are no moral obligations or refrains from making moral judgments: in any case, ‘immoralist’ should be taken in the sense given here.)

An impressive number of critics of emotivism encountered in the literature and in conversation have thought that there are obvious cases of immoralism (Brink 1989, chapter 3; Campbell 1993; Kupperman 1999, chapter 5; Mele 1996; Sinnott-Armstrong 1993; Smith 1992; Smith 1998; Simpson 1999; and Stocker 1979). Consequently, prominent figures in metaethical texts are Satan, who wants to do what is wrong just because it...
is wrong; the practically irrational person, either in the guise of the morally weak, who knows what is right but is overcome by desire and acts against his conviction, or in the guise of the depressed individual, who knows that he ought to help, say, but lacks all motivation; and the occasional philosopher thinking that he ought to report his gambling income to the IRS but feeling no inclination to do so. The simple idea is that, since such characters are real or at any rate possible, and since they would be ruled out by the emotivist identification of moral opinions and motivational states, emotivism must be false.

Emotivism is typically understood as the combination of two ideas: internalism and noncognitivism. Internalism says that moral opinions essentially involve some kind of motivational state, that their psychological function is to bring about or prevent action of the kind with which they are concerned. Noncognitivism denies that moral opinions are representations of moral reality: although they will tend to correspond to various features of the actions with which they are concerned, this is not their primary function. Quite obviously, immoralism is primarily a concern for the internalist part of emotivism, and there are many other views that involve internalism. But exactly how immoralism is such a concern and what kinds of defense are available will depend on the exact form of internalism. What I hope to do in this paper is to give an outline of an internalism that is a natural development of the emotivist theory proposed by Charles L. Stevenson and to show how it can account for at least some typical cases of apparent immoralism and make it plausible that other cases can be handled along similar lines. Seemingly clear cases of immoralism in the real world do not stand up to scrutiny, and methodological considerations show that very little weight should be given to imagined cases. Emotivism thus remains a live alternative in metaethics. Or so I will argue.

1. Emotivism

Emotivism, as it was formulated by Stevenson, claims that the function of moral terms is to express attitudes of some kind and cause similar states of mind in the hearer when put in the suitable sentential and conversational context (Stevenson 1937, 1944). (I take some liberties with Stevenson’s account, talking about functions rather than causal dispositions, for example.) The function of sentences such as “it would be morally wrong to privatize public land” or “it would be wrong to ban the private ownership of handguns” is to express a negative attitude to acts of, respectively, privatizing public land and banning the private ownership of handguns and to produce a similar attitude in the audience.

Given this account of the meaning of moral or ethical terms, and given the assumption that the function of (unembedded,
How Emotivism Survives Immoralists, Irrationality, and Depression

categorical) ethical sentences is to convey moral opinions, the following tentative account of one class of moral opinions seems reasonable:

An Emotivist Suggestion: To have the opinion that it would be morally wrong to perform an act of kind \(K\) is to have a negative attitude toward acts of kind \(K\).

I take it that this is part of the core of a viable emotivism. It is a central claim that figures in emotivist explanations of various features of moral thought and thinking, and emotivism is (insofar as we follow Stevenson) above all an explanatory theory, hoping to give us a deeper understanding of the causal structure of moral language and our moral lives. Other parts of that core would say something about our thoughts about what is morally admirable and what is morally permissible and probably extend to a fairly wide range of normative or evaluative opinions, including, perhaps, certain opinions about what is rational, as in Gibbard 1990.

The complications involved in a full-fledged emotivism are not needed for the points I hope to make against the immoralist argument. However, unless we have a clearer notion of the kind of attitude mentioned in An Emotivist Suggestion, there is very little we can say about the immoralist challenge. Clarifying this notion, we will have to make some theoretical commitments, and my characterization of the relevant attitude will place emotivism within a general functionalistic account of psychological states. Functionalism is of course not an uncontroversial approach in the philosophy of psychology—although perhaps the most influential and promising one—and if functionalism doesn’t work, neither do some of the arguments I will give later on.

First, let me introduce a technical term for the relevant kind of attitude: ‘attitude’ is itself too vague and general to capture what emotivism should be concerned with, and while ‘desire’ is almost right, it has phenomenological connotations that make emotivism untenable—to think that some act is wrong is not to have any particular feeling. Instead, I will define the notion of a moral optation and restate the identification using that notion.

I take an optation for acts of kind \(K\) not to be performed to be a psychological state the function of which is to prevent acts of kind \(K\) from being performed, through preventing the formation of intentions involving the realization of acts of kind \(K\). Optations so conceived are ‘propositional attitudes’, mechanisms operating on representations of actions so as to cause or prevent the realization of the action represented. Exactly how to ascribe functions is not something I will go into here, but I will suppose the following: the function of a thing is something it does under ‘normal’ or ‘ideal’ circum-
stances. That something fails to produce a certain effect under certain circumstances is entirely compatible with its having as its function to produce that effect, and even with its having as its function to produce that effect under those very circumstances. In particular, optations can fail to produce their effect because the opportunity never comes up, because we lack the means, because we lack knowledge of the means, or because there is a conflict between optations; the list of possible circumstances under which an optation can fail to perform its function can indeed be made very long. Similarly, a heart and a door lock can retain their functions of, respectively, pumping blood and of keeping people without the right key from opening a door, even when they are malfunctioning or just not being used. I will also suppose that the notion of ‘normal’ or ‘ideal’ circumstances can be spelled out in non-normative terms. (Perhaps the most detailed attempt to give precise, naturalistic, non-normative definitions of function and normality is found in Millikan 1984.)

Of course, given this definition of ‘optation’, not any optation is of the kind that emotivism should identify with opinions to the effect that some kind of act is morally wrong. I don’t want to eat right now, and this is presumably a matter of my having an optation for me not to eat right now, but certainly not an optation that could be identified as my opinion that it is morally wrong for me to eat right now. Now, to characterize moral optations with enough precision is a tricky affair as pointed out by Alston (1968), but I think the following connection to moral emotions such as guilt and indignation takes us a long way:

A moral optation for acts of a certain kind not to be performed is a psychological state the function of which is to prevent such acts, at least partly through eliciting guilt in the agent contemplating performing such acts and indignation toward others performing or contemplating performing such acts. (See Björnsson 1998 as well as Gibbard 1990 on morality in the narrow sense.)

Obviously, this definition assumes that guilt and indignation can be understood without reference to moral opinions. Admittedly, that is a controversial assumption. It might seem that guilt essentially involves the thought that one has done something wrong in much the way that hope involves the thought that what is hoped for is possible but not entirely within one’s power. But on the other hand, it seems that we can feel guilt even when we think that we have done nothing wrong. In any case, since my prime concern here is the immoralist argument and not the difficulty of characterizing moral optations, and since the latter problem has been addressed in Björnsson 1998 and Gibbard 1990, I will assume that the above
definition is on the right track and propose the following more definite formulation of the tentative identification given in An Emotivist Suggestion:

The Emotivist Identification: To have the opinion that it would be morally wrong to perform an act of kind $K$ is to be in a psychological state the function of which is to prevent such acts, at least partly through eliciting guilt in the agent contemplating performing such acts and indignation toward others performing or contemplating performing such acts.

The Emotivist Identification could be supplemented with the explicit noncognitivist claim that moral opinions do not represent moral facts in any interesting sense or that the representational content of moral opinions is secondary to their optational function. But none of my arguments here will depend on any such supplementation.

2. Methodological Considerations

The Emotivist Identification is meant to be methodologically on a par with other theoretical identifications, like the identification of light and photons/electromagnetic waves of a certain kind and the identification of instances of water and stuff constituted by $\text{H}_2\text{O}$ molecules. Note two features of such identifications: (1) Their acceptability is a matter of the theoretical characterization being at the core of our best general explanation of the causal upshots and effects of the phenomenon in question. Our knowledge of theoretical identifications is thus an a posteriori matter. So, I suggest, is our knowledge of whether emotivism is true. (2) They are necessarily true statements that support counterfactuals—insofar as they are true. Similarly for The Emotivist Identification, insofar as it is a correct identification, it is a necessary truth. Or more cautiously, since modal judgments and counterfactuals are sensitive to context, within scientific discourse or from a scientific point of view, theoretical identifications are necessarily true statements. Presumably, they do not bar fiction from meaningful talk of light with different physical nature or water with different chemical nature. Nor should emotivism imply that fictional accounts of people having moral opinions without corresponding moral optations are meaningless or mistaken.

Although both points can be questioned, I will just assume that they are correct. Enough philosophers seem to agree that they are correct for the following to be an interesting claim: if (1) and (2) are correct, emotivism is not threatened by arguments from immoralism. That is what I will argue.

When assessing theoretical identifications, appeals to what we can imagine, unconstrained by what our best explanatory
theories say, show very little. I can imagine that Jesus makes wine out of water without added ingredients, that the physical nature of light is changed by Satanic intervention or pure chance, and that Mark Twain isn’t Samuel Clemens but, rather, is Samuel’s unknown twin. If this is enough to undermine the identification from physics of light and photons/electromagnetic waves, the identification from chemistry of being a sample of water and being a sample consisting (almost exclusively) of H₂O molecules, or the identification from biography of Samuel Clemens and Mark Twain, then *identifications* are much less interesting than they seemed to be: all emotivists should strive for in the case of moral opinions is what we think chemistry has given us in the case of water. On the other hand—and this is what I suggest—we might deny that our imagination is a reliable guide to theoretical identifications: intuitions about imagined cases might explicate intuitive knowledge about moral opinions and help armchair reflection yield fruit, but they might equally well explicate mistaken assumptions.

The latter seems especially plausible when competent people have different intuitions, as is the case with imagined immoralists: philosophers who think that moral opinions have a necessary tie to motivation have denied that the characters we imagine when we imagine immoralists do have the moral opinions in question or insisted that they must have some motivation: what we conceive of are not really immoralists after all (Dancy 1993, 4–6; Gibbard 1993, 318–9). Such a response is not by itself enough to take the sting out of the immoralist, as David O. Brink insists in his version of the immoralist argument (e.g., Brink 1989, 46–7), but it does call for serious reflection on the reliability and relevance of our modal intuitions.

The test of theoretical identifications, then, shouldn’t be imagination but explanatory power. The reason we should accept what chemists tell us about the nature of water is that the assumption that instances of water consists of H₂O molecules is central to good explanations of how water interacts with or is created in interactions between other substances. Similarly, emotivism provides a reasonable theory of the psychological nature of moral opinions insofar as *The Emotivist Identification* is central to good explanations of the causal upshots and behaviors of moral opinions.

So, in assessing emotivism, we should attend to its explanations of the causes and effects of *real* moral opinions as well as to the explanations offered by rival theories. Moreover, the occasional case apparently involving a moral opinion but no moral optation need not be a problem. Chemists claiming that sugar is a substance constituted by soluble crystalline carbohydrates should not be shaken by the fact that aspartame can look and taste sugary although it is of a different chemical constitution. Rather, the correct conclusion for the chemist is
that aspartame is a different substance. Similarly, the occasional apparent case of immoralism should not trouble emotivists if their theory provides good explanations of what goes on in most typical cases and if there is no description of moral opinions that has similar explanatory power but captures more cases. (It is also possible that no single characterization will do fundamental explanatory work for the great majority of instances of a phenomenon: we might find, as in the case of jadeite and nephrite, that what has fallen under one label, 'jade' needs not one, but two alternative theoretical characterizations. But I see no need for that move in the case of the moral opinions with which The Emotivist Identification is concerned.)

Of course, this kind of assessment relies on the possibility of finding clear cases of moral opinions, and that might be more difficult than finding good samples of water or sugar. Still, I think that we can agree on numerous cases where moral opinions obtain and cases where they do not, and we can learn about what the causes of such opinions are and what difference they make to the behavior of people having them. As a matter of fact, I think that we know quite a lot about how moral opinions interact with other psychological features under normal circumstances. The task of the philosopher or moral psychologist is to suggest causal structures that could explain such interactions and to say something interesting about these structures, for example explaining why they are there to be found at all and along what general principles moral opinions interact with other parts of the structures that are our psyche. Given the complexity of the human biological system, we can at most hope to find principles that hold ceteris paribus or under some especially interesting conditions. But this is a problem for any study of complex phenomena and not cause for abandoning further inquiry.

What I propose to do here, then, is to look at real cases of what seems to be immoralism of fairly familiar and widespread kinds, such that if these kinds of apparent immoralism turned out to be not only apparent but real, this would suggest that the causes and effects of moral opinions are generally best understood from a different theoretical identification. But first, let me say what assumptions I will make about the capacity of The Emotivist Identification to explain other phenomena associated with moral opinions.

3. The (Presumed) Evidence for Emotivism

Since this paper is exclusively concerned with the argument from immoralism, my claim will be that if emotivism does well in other areas, the argument from immoralism fails. I will therefore unabashedly assume that emotivism is an explanatory success elsewhere.
For some areas, this assumption is very plausible. It is hard to think about our moral lives without noting that we have a strong tendency to feel guilt when we think that we have done something wrong and a tendency to feel indignation when we think that someone else has done something wrong. And it is hard to ignore the fact that guilt and the indignation of others keep us from acting in ways eliciting such feelings. Moreover, it is generally agreed that motivation to act and emotions like guilt and indignation are important signs of moral opinions and that the lack of motivation and emotional reactions are important signs of lacking moral conviction (e.g., Stevenson 1944, 16–17; Alston 1968, 10; Smith 1972; Dancy 1993, 4; Smith 1994, 6). On emotivism, that is just what we should expect, and this gives prima facie support to the theory.

However, I will also assume that emotivism can account for phenomena that have been seen as beyond its explanatory powers. For example, I will assume that emotivism can explain the role of moral opinions and moral statements in inferences and the fact that moral statements seem to figure in logical relations in the same way as paradigmatically non-emotive statements. Also, I will assume that The Emotivist Identification provides good ground for an explanation of our intuitions about moral mistakes and disagreement and that it can explain how we can feel guilt even though we think that we have done nothing wrong. These are controversial assumptions, to be sure, but since I have argued for their correctness at length in Björnsson 1998 and 2001, I will just take them for granted and try to make plausible that if they are correct, the argument from immoralism has little power.

If all this would work—if the causal role of our moral opinions is well explained from the assumption that they are moral optations—we would have a nice little argument for emotivism. On almost any view, the moral opinions of most people coincide with such things as moral optations: motivational states tightly connected to emotions like guilt and indignation. And if The Emotivist Identification would turn out to be an explanatory success, there would be no reason to postulate some further entity—some non-optational moral belief—doing the relevant causal work. (If moral optations can have truth-conditions—if they might be ‘besires’ as Altham 1986 put it—emotivism is compatible with an internalist form of cognitivism.) But all this depends on there being good answers to the argument from immoralism.

4. Real Immoralism: A Closer Look

In line with the methodological considerations in section 2, I will concentrate on kinds of apparent immoralism that are ordinary enough to spell trouble for The Emotivist Identification should they turn out to involve no optation and try to explain
how they come about and interact with other phenomena, ultimately deciding whether they do indeed involve moral optations. I will discuss four categories.

(1) Start with moral weakness. Clearly there are cases where we think that it would be wrong to do something but go ahead nonetheless: this is something most of us have some experience of, if not recently. What such cases show, however, is simply that moral optations sometimes fail to perform their proper function: they fail to produce or prevent the relevant actions. In doing this, moral optations are similar to a wide range of desires, as evidenced by the phenomenon of conflicting desires, where a subset of our desires takes control over our behavioral resources and determines action in spite of desires to the contrary. This can hardly show that emotivism is wrong, as emotivists have repeatedly argued.

(2) In cases of moral weakness and weakness of will, the overridden motivational states typically have effects such as regret, guilt, shame, ambiguous motions or facial expressions, and so forth. That is what makes talk about weakness so apposite: there seems to be a struggle going on, where the immoral or irrational side eventually gets the better of the moral or rational. However, cases of depression and listlessness do not seem to involve a motivational struggle and might therefore provide clearer instances of immoralism. According to arguments pointing to such cases, the depressed or exhausted individual can have the opinion that it would be wrong of her not to do something, without therefore having the least motivation to actually do it. In giving extra weight to this assumption, Michael Smith reports that what is really bad about depression is that one knows that one would have had the motivation to do the right thing, had one not been depressed (Smith 1998, 149–74).

I agree with Smith’s observation, at least as it concerns the mild kind of depression or listlessness of which I myself have some experience. But what do observations of depression show? Rather than looking at chronic depression, consider something that most of us will have an easier time thinking about, namely temporary states of listlessness, due to temporary exhaustion, sleep deprivation, sickness, or personal loss. Suppose I find myself in such a state, remembering that I had promised to meet a friend downtown to help with a few things. I think to myself that it would be wrong not to go, but that thought does not engage, somehow, and I find myself lying down on the sofa instead of putting on shoes. Four complementing considerations suggest that emotivism survives examples like this.

(i) While it is plausible to say that the listless has moral opinions but lacks moral motivation, it often seems equally plausible to say that she wants to go help her friend, say, but lacks the
energy or willpower. But it is plausible that the psychological function of wanting to do something is exactly to make one do it. If so, the listless is in a motivational state of the relevant kind—a state the function of which is to make her go help her friend—and we have no clear counterexample to emotivism.

(ii) Emotivism claims that moral opinions are states the function of which is to affect action in a certain way. But introspection is obviously not an infallible guide to the functions or the presence of states that govern our behavior and thinking, especially not under abnormal psychological conditions. Hence, even if the listless doesn’t think she is in a certain motivational state, she might well be.

(iii) Why does the listless seem to have a moral opinion? There is perhaps a certain readiness to utter certain sentences and a tendency to give reasons for her opinion, and perhaps the thought that “It would be wrong not to go” passes through the conscious mind without any feeling of insincerity (apart from that having to do with the lack of drive). But the discussion in this paper assumes that emotivism can explain why moral optations tend to produce utterances, conscious thoughts, and moral reasoning. If this is right, these phenomena are equally signs of moral optations. Hence, there is evidence that the listless does have moral optations. Moreover, there is nothing strange about moral optations manifesting these dispositions under conditions of listlessness while losing their power to influence action or some of their motivational phenomenology: different dispositions are in general manifested under different circumstances.

(iv) This brings up the final point. A moral optation is a state that affects action in specific ways under ‘normal’ or ‘ideal’ conditions for affecting action, and listlessness is hardly such a condition. Listlessness and depression are general motivational disorders that lower energy levels and one’s motivation to do anything, including mischief. This generality suggests that it is not particular optations that are lacking in the listless or depressed, but general means for their enactment. (Medical science suggests that low levels of dopamine and serotonin in the brain play a crucial role.) Moreover, this would explain the following commonsensical observation: when cured from our listlessness, we usually find ourselves with much the same motivation and range of desires as we had before the affliction. So, even assuming emotivism, it is no wonder that the moral opinions of the listless come without the specifically motivational phenomenology. (As suggested by Aristotle 1980, Book 7, those “asleep, drunk or mad” can also be said to be under a general motivational
How Emotivism Survives Immoralists, Irrationality, and Depression

deficiency rather than lacking particular motivational states. Gibbard (1993, 318) might be advocating a similar response.

It seems, then, that emotivists should not worry about these kinds of practical irrationality or deficiency. And various pathological cases can be handled along similar lines. Take, for example, the cases of brain damage described in Damasio 1994. One of Damasio’s patients is able to reason well about what would be the morally correct action in various difficult cases, as well as about what would be reasonable decisions in his own life. However, his actions always seemed to be related to more primitive impulses, leading to disastrous results. The interpretation that I suggest for a case like that is that this person has the moral optations that a normal person has, but that they fail to affect action due to missing parts of the brain.

(3) There are, of course, common forms of apparent immoralism that do not seem to involve practical irrationality. Suppose that someone thinks that it would be wrong not to report his betting income to the IRS but, nevertheless, feels no inclination to report it. He assumes that he lacks motivational states the function of which is to make him report that income and concludes that emotivism must be false (Sinnott-Armstrong 1993, 302). But why should he conclude from not feeling inclination that there is no moral optation? We have already seen ways in which optations can fail to be expressed, and this case gives us reason to look at more such ways. Since I have heard similar arguments from people with whom I am personally somewhat acquainted, I will draw from this knowledge, though I will stick to the example above for purposes of illustration. Typically, the following seem to be true about this kind of apparent immoralist:

(i) The person’s moral optations are in many other cases clearly aligned with his moral opinions. In particular, they are so aligned when concerned with acts similar to not reporting the betting income, such as a wealthy person’s act of not reporting large parts of a high salary.

Now, our optations are not randomly distributed over various actions, but sensitive to features of those actions, features that can be exemplified by other actions and eliciting similar optations in similar cases. Given (i), there is a reason to expect a moral optation concerning the IRS report, an optation elicited by the same features of failing to report the betting income that trigger the optation concerning the wealthy person’s not reporting parts of his salary. But if this is indeed the case, why is there no felt inclination to act accordingly in this case? Unlike the case of depression, there is no general motivational deficit, so the explanation would have to lie
elsewhere. What I suggest is doing important work here is that:

(ii) Reporting the betting income involves a cost on the part of this person (filling in a complicated form, not being able to go on that vacation, or just losing a little money): the person has an optation not to report the betting income.

If there is a moral optation, there is also a conflict of optations, just as in cases of weakness of will. But why, then, is there no sense of struggle in this case? The answer, I think, is that the moral optation is silenced by the optation not to suffer the personal cost. Such silencing is not a strange phenomenon invoked ad hoc here: rather it is a basic and very useful feature of practical reasoning. When we are considering alternative routes of action and choose a route that we think best overall, that route nevertheless fails to realize some goal that another route would have realized. In difficult decisions—decisions where we are likely to talk about moral weakness or moral strength—we are painfully aware of this, but it takes place even when decisions are easy. Once such a decision is made, there is typically little point in spending cognitive resources on what isn’t but could have been achieved: that, I suggest, is the general reason why optations elicited by considerations not adhered to are silenced.

Why, then, are not all optations contrary to decisions silenced? Why do we experience weakness, regret, or discomfort with our decisions? The answer, I suggest, comes in two parts. The first is that silencing might be a bad thing when the disregarded optation is an important one, where it might be important to be on the lookout for options that involve its satisfaction. The other is that silencing might be a bad thing when disregarding an optation is likely to lead to adverse reactions such as disappointment, anger or indignation on part of others. But in cases like the IRS example, the following would typically be true:

(iii) Doing what one thinks is wrong is unlikely to lead to adverse reactions, and it is not something that one regards as very wrong.

This seems to be the case with the IRS example. Tax evasion is not generally considered a grave moral sin—in fact, the IRS is seen by many as an enemy—and a common thought is exactly that no one (identifiable) will be hurt. Hence, silencing could very much be expected in such cases. Of course, (i) does not give a conclusive reason that there is a moral optation that can be expected to have been silenced given (ii) and (iii); it only makes the existence of such an optation quite likely. But if we assume—
as I do in this paper—that emotivism is well supported prior to the analysis of immoralists, and if we hold in mind that knowledge of our own motivational states is limited, a good likelihood is all emotivism needs to retain credibility.

An interesting case somewhat similar to the IRS example is provided by brands of cynicism where someone is doing what she thinks is wrong because others are doing the same thing and because only one person doing what is right would fail to significantly help the situation. Disillusioned, she lives a life that she thinks is immoral. Again, she might not feel much regret, but that does not mean that she is missing the relevant options: silencing can be quite effective. So, emotivism can explain how cases like the IRS example is possible, and why there is a difference between moral opinions as to how obvious motivational signs are involved. Rather than being a problem for emotivism, the phenomenon gives it further support.

(4) Before turning from real cases of apparent immoralism to facts about what people are willing to say and think about fictional characters such as Satan, we can have a look at a real ‘satanic’ act, reported by St. Augustine in his Confessions. St. Augustine confesses that, in his youth, he and some friends once stole pears from a neighbor, just because it was wrong: they had no interest whatsoever in the pears and could get much tastier fruit along morally acceptable and less strenuous routes. Is this a problem for emotivism? I don’t think so. Kids—and adults—do all kinds of things just because they are (perceived as) dangerous or disgusting. Admittedly, such cases need not involve any clear feeling of practical conflict, but the feeling of danger involved is very plausibly seen as the expression of a state the function of which is to make one avoid the activity: the excitement and heightened sense of being alive results from mastering the dangerous situation and overcoming one’s own reluctance. To take an obvious example, people bungee-jump just because it is scary, because there is a reluctance to master. The same, I suggest, holds when people—like the young St. Augustine—do what they think is wrong “just because it is wrong”: it is because parts of them “scream out” against doing it that what they perceive as wrongdoing is exciting and tempting. So emotivism does not rule out this brand of satanic motivation, quite the opposite.

I can hardly claim to have shown that the above kinds of apparent immoralism involve moral options. Partly, this is because the evidence is anecdotal; partly it is because I have been relying on intuitive but yet to be systematically documented principles of psychology; and partly because the relevant notion of function remains to be given a precise definition. Regarding all this, I claim bad company: the immoralist arguments against emotivism have equally—or to a greater extent—relied on anecdote and folk psychology, and my
goal here has been to undermine those arguments, not to show that emotivism is correct. Moreover, science rarely starts out with precise definitions, but adds refinement as evidence demands it. Enough refinement has been supplied to undermine the arguments from immoralism discussed so far, and that has been my primary ambition.

5. The Emotivist Identification and the Commonsense Notion of Immoralism

Consider an argument taken from what is probably the most nuanced treatment of immoralism in the metaethical literature. I quote from Ronald D. Milo's *Immorality*:

> we think that the most evil or reprehensible kind of wrongdoing consists in willingly and intentionally doing something one believes to be morally wrong ... because one simply does not care that it is morally wrong.... Any analysis of the nature of moral beliefs that does not allow for even the possibility of these forms of immoral behavior is, it seems to me, defective. (Milo 1984, 253)

Similar arguments are found in Brink 1989 (27) and Copp 1995 (16). I have already argued that our intuitions about the possibility of immoralism should be given little weight when assessing emotivism, but isn’t the world plentiful of reprehensible wrongdoing of just the kind Milo is pointing to? In any case, this is certainly what people seem to think, and so emotivism seems to imply that people are in massive error. Is that really plausible?

Here, one could perhaps argue, as Nietzsche famously did, that everyday psychological belief—especially when concerned with *evil*—is replete with misunderstanding boosted by Christian mythology. (See for example Nietzsche 1994.) However, we can find less intricate ways to explain the commonsense conception of reprehensible wrongdoing without *ad hoc* assumptions, and what can be explained on a given theory can hardly tell against it, quite the contrary. But let me start with some considerations that should free us, first of the idea that the common sense understanding of immorality must be veridical, and then of the idea that it is necessarily in conflict with emotivism.

To begin with, although *depictions* of the kind of wrongdoing with which Milo is concerned satiates a plethora of soap operas, the real thing is not something we encountered daily. Now, thoughts about cases of which we have very little experience tend to lack reliability: remember how contrary to general expectations the data from Stanley Milgram’s experiments on obedience were, and consider how it surprises many people that for some of the most atrocious evils such as genocide, the perpetrators have thought that what they did wasn’t wrong, but
rather morally compulsory. (See for example Staub 1989.) Hence, the thoughts Milo alludes to might tell us little about reality.

Moreover, it is not clear that what common sense has in mind when thinking of the most reprehensible wrongdoing involves a person lacking moral optation to abstain from what he knows is wrong, where moral optations are understood in the sense defined above. I have already argued that the presence of such optations is compatible with little or no overt behavior and little or no feeling of motivation. These have not been obvious points even in a philosophical context, and so common sense should not be expected to have grasped them adequately. And even if this is so, it is not clear that it is a complete lack of moral optation rather than a lack of motivationally effective moral optation that is central to the commonsense conception of reprehensible wrongdoing.

So, it is neither clear that common sense has a veridical grasp of reprehensible wrongdoing nor that whatever vague grasp it has is in conflict with emotivism. But furthermore, emotivism can explain why we would extend our everyday grasp of moral opinions to cases where few of the standard motivational signs obtain.

The first assumption in this explanation is that having a moral optation has a fairly wide but reasonably systematic range of psychological and behavioral effects, including a tendency to adduce supporting reasons and instigate inferential moves: this I have taken for granted. It would be plausible to expect our thinking about moral opinions to be adapted to this systematic range of effects, so that we expect these effects when encountering moral opinions and identify moral opinions by these effects.

Given this, what would happen if we came across a person—real or fictional—that displayed all but a relatively well-defined subset of these psychological and behavioral patterns? How would we describe such a person? That would depend on whether we had a specific expression signifying the specific state of mind giving rise to that limited set of effects—in which case we would use that expression—but suppose we had no such expression. Then, I suggest, we would use our ordinary expression for moral opinions and note in what ways the usual assumptions fail to hold: the deviant case would be described on the pattern of “P thinks that K is morally wrong, but X is the case,” where X signifies what is different from standard cases.

Apply this to some purported or fictional immoralist: “He knows that torturing political prisoners is wrong but lacks the motivation to abstain from such torture.” Here, the audience could use their everyday conception of moral opinions; deduct the motivational part but keep the rest of the ordinary assumptions more or less intact; conclude that this disagreeable
character knows that the torture is extremely painful for the victim, that it often is psychologically devastating, and that people in general think that it is wrong and are emotionally agitated by it; and further conclude that, partly because of such knowledge, he has the capacity to verbally and otherwise behave somewhat like normal people thinking that torturing political prisoners is morally wrong. The usual descriptions of immoralists are thus to be expected, given emotivism, and given that we lack a special expression for the state in which apparent immoralists find themselves. The same phenomenon of using an expression denoting a substance, \( S \), to convey thoughts about objects that share many but not all of the typical properties of \( S \) also explains the—at least apparent—intelligibility of talk about Jesus' changing the chemical composition of water and about Mark Twain's being Samuel Clemens's unknown twin. Moreover, given the questionable veridicality of our everyday understanding of reprehensible wrongdoing and the general difficulty of understanding moral opinions that is displayed by ongoing philosophical controversy, it is quite likely that we should be unaware that our description of the wrongdoer, taken 'literally' or 'scientifically', describes the impossible.

“But isn't it just very plausible that some real, non-depressed, psychopaths know that what they do is wrong, without therefore being the least motivated to refrain?” If this thought goes through the reader's mind at this point, I can only ask her to look through this section once more, considering (a) whether the thought really is in conflict with the present form of emotivism and (b) how likely it is to carry information about how to characterize our moral opinions so as to provide information central to our best explanations of their causes and effects.

6. Concluding Remarks

But maybe the argument of this paper shows too much. Consider Joe, claiming that all birds fly, and dismissing penguins as not really being birds just because of their lack of aviation: can't he use the same arguments to dismiss objections to his claim? Not so. This case is different from the emotivist treatment of putative immoralists in that Joe goes against fundamental organizing principles of biological taxonomy that focus on internal structure and evolutionary background. (This focus is not arbitrary but chosen for structuring information in ways conducive to biological inquiry and explanation.) Unfortunately for Joe, penguins are birds on both counts. By contrast, it is unclear how emotivists flout any such principles. Quite the contrary, one popular and well-advocated taxonomic principle in philosophical psychology (and, of course, in physiology) is to categorize in terms of function, and this is exactly what our version of emotivism does, basing the ascrip-
tion of function on observation of interesting ways in which moral opinions normally interact with other states of mind and with other people through emotional and linguistic expression. A more telling example is this: if Joe found or dreamt up an animal very much looking and behaving like a penguin but nourishing its young with milk and stemming from other mammals, biologists would neither say that it is a bird nor that it is a penguin. The temptation to talk about the creature as ‘Joe’s Penguin’ would be neither here nor there.

The literature on emotivism contains too many varieties of apparent immoralism to examine in one paper, showing for each type how it can be accommodated within an emotivist framework. What I hope to have established, however, is the general difficulty of undermining emotivism or internalism with reference to immoralism and an awareness that if we are interested in the nature of moral opinions (rather than in our conceptions or beliefs thereof), fantasy and intuitions have little force. If anything, the cases we have looked at have shown the resources of emotivism and hinted at ways in which other forms of internalism might be defended.

Does this mean that there is no straightforward way to refute emotivism? Certainly not, if it is false: the argument here has relied on the assumption that The Emotivist Identification is central to explanations of the causes and effects of moral opinions in ‘non-immoralist’ cases, and this assumption is of course far from trivial. There are the prima facie difficulties of explaining moral reasoning and the structure of moral disagreement and of accounting for cases where we seem to have moral optations but nevertheless lack corresponding moral opinions, to mention a few. And if the presence of moral optations fails to do the relevant explanatory work, we might have reason to reject emotivism even if all clear real cases of moral opinions come with the right kind of motivation. But these are large issues for other occasions.

Note
This paper has benefited from being discussed at the Higher Seminar in Practical Philosophy at Stockholm University and at a Brown Bag meeting at the Philosophy Department at the University of Connecticut. Detailed and helpful comments have also been given by Paul Bloomfield, Joel Kupperman, William G. Lycan, and Folke Tersman. Joe’s Penguin was conceived as an obstacle for the emotivism defended here by an anonymous referee for Philosophy and Phenomenological Research. The paper was written while on a grant from STINT, The Swedish Foundation for International Cooperation in Research and Higher Education.

References
Gunnar Björnsson


