Conceptualizing Metaethics

A Commentary on Prinz

Yann Wilhelm

In this commentary on Prinz's "Naturalizing Metaethics" I shall first look briefly at his methodological assumptions. I will argue that Prinz's approach is more radical and less conciliatory between analytical and empirical approaches than it seems from his own description. In the second part of my commentary, I shall look at one possible objection to Prinz's sentimentalism: the evidence he presents does not provide the needed modal strength for sentimentalism. I shall present two example of this objection, and argue that Prinz's own depiction doesn't adequately represent it. I shall then use the helpful distinction offered by Jon Tresan between *de dicto*- and *de re*-internalism to analyze underlying problems in the objection. I will present another way of reacting to it, which I think fits nicely with Prinz's naturalized methodology. In the last part, I shall look at his critique of non-cognitivism. Prinz argues that non-cognitivism makes certain linguistic predictions that turn out to be wrong: if non-cognitivism were true we would expect our moral language to reflect this. I will argue that there are many forms of non-cognitivism that predict this surface grammar. The key idea is that non-cognitivism entails a pragmatic theory of moral language. I then offer a speculative explanation about why the moral language has its surface form. This speculation, I argue, has at least the same amount of plausibility as cognitivist theories. Furthermore, this possible explanation is open to empirical investigation. I agree with Prinz that, ultimately, metaethical theories should be tested against empirical evidence. Prinz presents conceptual and empirical work as mutually enhancing enterprises. My commentary is, I hope, a small contribution highlighting the conceptual side of the coin.

Keywords
Cognitivism | De dicto-internalism | De re-internalism | Metaethics | Methodological naturalism | Motivational internalism | Non-cognitivism | Sentimentalism

Commentator

Yann Wilhelm
ywilhelm@students.uni-mainz.de
Johannes Gutenberg-Universität
Mainz, Germany

Target Author

Jesse Prinz
jesse@subcortex.com
City University of New York
New York, NY, U.S.A.

1 Metaethics under empirical scrutiny

Prinz proposes to naturalize metaethics. Metaethics is traditionally regarded as a second-order discourse about ethics. Where normative ethics asks what is good and what is bad, what we should or shouldn't do, metaethics asks the question of what morality is itself (DeLapp 2011). Its subject is the ontology of moral properties, the semantics of moral discourse, the epistemic foundation of moral judgments and the psychology of moral opinions. These different aspects are highly interrelated—answers in one area influence questions asked in others.

There are many different ways to tackle the question of what morality itself actually is. Prinz characterizes metaethics as being concerned with the foundations of moral judgments (Prinz this collection, p. 1). This is his starting point, which shapes his decision tree. He acknowledges that one could arrange the tree in different ways, depending on which aspect one wants to pull into focus.

Prinz's primary goal is to show that every question in the decision tree is empirically tractable (this collection, p. 1). This is his *methodological naturalism* (p. 2).[1] He argues that we

1 He contrasts this with *metaphysical naturalism* and *semantic naturalism*. The former says that everything there is belongs to the natural world. The latter tries to reduce concepts from various domains in terms that are more likely to be naturalized in the metaphysical sense.

should study the domain of metaethics empirically. He wants to test "[...] theories derived from philosophical reflection against the tribunal of empirical evidence" (Prinz this collection, p. 5).

Metaethics, according to him, is not the sole matter of armchair reflection. This seems natural when we characterize metaethics as the question of what morality itself is. But that goes against the view that metaethics—or philosophy in general—is not concerned with what actually is the case, but with what *must* be the case. What are the *necessary* conditions of morality? On this view, metaethics is concerned with statements that hold *a priori*. Most of the time this means deriving knowledge from reflection upon the meaning of our concepts. This method of *conceptual analysis* had been at the core of philosophy since the *analytic turn* (Prinz this collection, p. 3).

Against this turn Prinz sets the *empirical turn* (this collection, p. 3). He describes this development as an enrichment of the philosopher's tool box. Where conceptual considerations help us to formulate theories and flesh out the differences between different views, empirical methods confirm the theories derived from this work. The former pose questions and formulates possible answers; the latter test those answers. Prinz emphasizes that empirical and traditional approaches are not opposed to one another (this collection, p. 5). Rather, they complement each other. They're more like opposing points on a continuum of methods for exploring the world.

It is important to see that this view is not as conciliatory between traditional analytic philosophy and empirical philosophy as it might seem. It does not leave room for *a priori* armchair reflection. In fact, Prinz even regards conceptual analysis as an empirical task: "[A]rmchair conceptual analysis can be characterized as an introspective memory retrieval process. As such, it can be regarded as a form of observation" (2008, p. 191).

When Prinz speaks of "traditional methods", he does not include conceptual analysis as an *a priori* enterprise. Rather, he is referring to various tools, for example formal semantics or logic, which help us articulate theories. They are tools for exploring the natural world, from which we gain knowledge only through experience. Prinz is a radical empiricist at heart.

An empirical scientist could ask: "What differentiates this from my own work?" For she, too, reflects upon different theories, how they relate to each other, formulates questions, and so on. This is an important part of scientific, empirical work. I think Prinz would agree. An important upshot of his naturalized philosophy is that there are no clear-cut borders between philosophy and psychology (Prinz 2008, pp. 204–206). They are different disciplines not because of their different subject areas or methods but for pragmatic reasons. They are different *academic* disciplines, shaped by sociological and historical processes. The borders between the different disciplines become blurred in the empirical turn. According to Prinz, this is a good thing.

I think this the real strength of Prinz's approach. Arguably many disciplines are divided largely by pragmatical differences, like education and academic organization. Instead of demarcating different approaches, instead of drawing sharp lines between them, Prinz proposes that we unite them in the search for explanations of the natural world.

Prinz's target article is a very good example of this approach. Here I want to make a few remarks in the spirit of Prinz's own methodology. In the next section I will focus on a specific objection against Prinz's answers to the first question in the decision tree. I think that it can clarify some consequences of his methodological naturalism for metaethics.

2 Internalism and modal strength

In this section I discuss Prinz's answer to a potential objection to his sentimentalism, namely, that the evidence lacks *modal strength*. In fact, objections of this kind have already been raised against Prinz's and other naturalistic metaethical theories already. I shall first argue that his answer doesn't get to the heart of the objection. Second, I propose a way in which Prinz can and should answer it. To do this I shall present two instances where this objection has been made. A helpful distinction by Jon

Tresan will then show that there are actually two kinds of internalist theses at play here. Only one of these is really relevant for Prinz's naturalized metaethics, I shall argue. The objection then loses its force in light of Prinz's project of a naturalistic methodology. The following reasoning can also be seen as a small case study in recent (naturalized) metaethics.

The first question in Prinz's decision tree is whether moral judgments are essentially affect-laden or not. This is Prinz's take on the internalist-externalist debate.[2] This debate is a classical debate in metaethics that can be traced back to the British moralists (Darwall 1995). It concerns the question of whether *motivation* is *internal* or *external* to moral judgments. Do moral judgments necessarily involve motivation to act accordingly? Or does the motivation come from a desire external to them (e.g., the desire to be a good person)?[3]

Prinz advocates a position that he calls sentimentalism:

> Sentimentalism $=_{\text{Df}}$ Moral Judgments essentially involve affective states, such as emotions, in one of two ways: such states as constituent parts of moral judgments (traditional sentimentalism); or moral judgments are judgments about the appropriateness of such states (neo-sentimentalism). (Prinz this collection, p. 6)

The evidence for a link between moral judgments and emotions is overwhelming (Prinz this collection, p. 10). But is it enough to warrant a stronger relation than mere accompaniment? Even if we grant Prinz the interpretation that affective states are not only mere *consequences* of moral judgments, could we not still question whether they are essential components of moral judgments? The objection is this: the empirical evidence lacks *modal strength* to support sentimentalism. Even if all our ordinary moral judgments are based on emotions, it could still be *possible* to judge dispassionately (Prinz this collection, p. 13). Therefore the evidence doesn't support sentimentalism.

Prinz answers that the empirical evidence gives us enough reason to infer that we *cannot* make moral judgments without emotions: "Every study suggests that emotions arise when we make moral judgments. All evidence also suggests that when emotions are eliminated, judgments subside as well" (Prinz this collection, p. 13).

According to Prinz, the theory that emotions are essential components of moral judgments explains the total pattern of data better than its rivals (this collection, p. 14). Furthermore, he argues that the sentimentalist can accept psychologically exotic cases, in which the connection between moral judgments and emotions doesn't occur, which conform rival theories.

This answer, I argue, misses the real core of the objection. Prinz confronts it upfront and just states what it questions. He puts the objection in the following way:

> The evidence shows that emotions are often consulted when making moral judgments, but this leaves open the possibility that we might also make moral judgments dispassionately under circumstances that have not yet been empirically explored. (Prinz this collection, p. 13)

But this does not represent the objection adequately. The objection doesn't rest on possible, not-yet-found empirical evidence against sentimentalism. Rather, it rests on opposing ideas about what kind of modal strength claims about the relation between moral judgments and emotions should possess. At the heart of this objection there is no disagreement about the empirical evidence, but an opposition in the underlying methodology.

Adina Roskies, for example, accepts that "[…] those [brain] areas involved in moral judgments normally send their output to areas involved in affect, resulting in motives

2 Although he doesn't explicitly put it like this, I think it's safe to frame it in this way. The option that denies affect-ladeness is called "externalist moral realism", and he states in various places that emotions are motivating or action-guiding (Prinz this collection, pp. 8, 11, 21). And one answer to the third question is a position called "internal realism". What I say about internalism in the following therefore applies equally to Prinz's sentimentalism. See also Prinz (2006), where he explicitly states motivational internalism.

3 See Björklund et al. (2012) for a short overview.

that in some instances cause us to act" (2008, p. 192).

But she thinks that this is not enough for internalism to be true.[4] In her view there is a connection between the cognitive and the affective system, but "this link is causal and thus contingent and not constitutive" (Roskies 2008, p. 192). In this sense the connection, according to her, is not necessary.

Antti Kauppinen sees the difference between internalism and externalism in a similar way. He depicts internalism as saying that there is a link between moral judgments and motivation that holds a priori and with conceptual necessity. externalism, in contrast, is the view that this link is contingent and a posteriori (Kauppinen 2008, p. 3). For Kauppinen, every internalist position then becomes an externalist position if it weakens the modality of the claim. When a metaethical account doesn't claim that the connection between moral judgments and motivation holds a priori and by necessity, it is an externalist account. No amount of empirical data can refute this criticism.

In Kauppinen's case the disagreement with Prinz about the underlying methodology is clear. He reacts to the proposal by Roskies, Prinz, and Alfred Mele (among others) that we clarify the debate empirically (Kauppinen 2008, 4). Because of his definitions of internalism and externalism as conceptual necessary claims he argues that "[...] findings in either actual or fictional experimental psychology or neuroscience have little relevance to the debate" (Kauppinen 2008, p. 4).

Kauppinen is opposed to methodological naturalism in philosophical moral psychology (2008, p. 4). That is why he would not be satisfied with Prinz's answer to this objection. Against him, Prinz would have to defend his metaethical naturalism. Interestingly enough, Roskies, on the other hand, thinks that we *can* clarify metaethical debates empirically.

In what follows I shall show how I think Prinz should meet this objection. Furthermore, I will argue that everyone who wants to apply empirical data to metaethical debates, such as, e.g., Adina Roskies, should side with Prinz on his methodological naturalism and accept internalism as a true a posteriori theory about moral judgments.

I will now present an analysis of the internalism–externalism debate offered by Jon Tresan that I think will be very helpful here (2009). He distinguishes different formulations of internalism along various dimensions. He claims that a very important distinction has been overlooked: most philosophers in the debate neglected the difference between the modality of the internalist claim and the stated relation between moral opinions and motivation. According to Tresan, there are two different kinds of necessity that can occur in such claims: wide-scope necessity, which operates over the entire proposition—*de dicto*—and narrow-scope necessity, which operates over the predicate—*de re* (Tresan 2009, p. 54). The first operates on the dimension of *Modality* and the second on the dimension of *Relation* (Tresan 2009, p. 55).

For example, the statement that parents have children can be formulated with both kinds of necessities:

> Necessarily, parents have children (*de dicto*).
>
> Parents have, necessarily, children (*de re*).

In the first case the proposition that parents have children is stated as holding necessarily. Parents have children, otherwise they would not be called parents. If someone has a child, she is a parent. But the second statement says that people who are parents have their kids necessarily. But this is obviously false. John and Mary don't have their children necessarily. They could easily never have had any children at all. True, they would not, then, be parents – but the fact that they are parents may have, initially, been quite accidental. We can easily see that there is a difference between *de dicto*- and *de re*-necessities because these two statements can have different truth-values at the same time.

4 Her critique is directed at internalism, not sentimentalism. But I regard both positions as similar enough to treat Roskies's critique as an argument against Prinz's sentimentalism (see also above). At the core of both positions is the connection between moral judgments and affective (motivational) states.

With this distinction at hand we can distinguish two different internalist theses: a strong Modality/weak Relation or *de dicto*-internalism, and a weak Modality/strong Relation or *de re*-internalism. The former states that, with necessity, there is a connection between moral judgments and motivation. The latter says that there is a necessary connection between these two things.

Tresan uses this distinction to argue that something has gone fundamentally wrong in the internalism–externalism debate. The neglect of the two features has led to the *internalist fallacy*: the strength in Modality of an internalist claim was taken to be strength in Relation, which led to an overestimation of the epistemic value of the claim (Tresan 2009, p. 55). The classical debate stated the connection between moral judgments and motivation in terms of conceptual necessity (a *de dicto*-internalism) (see Roskies's and Kauppinen's accounts above). Arguments for this claim were supposed to evoke the intuition that no one can make a moral judgment without being motivated to act. If we have such intuitions, the arguments go, the connection is a conceptual necessity. Likewise, arguments against this internalist claim consisted in thought experiments that were supposed to evoke contrary intuitions.

From Tresan's distinction follows that claims with *de dicto* necessity are claims about our concepts and not about the subject matter (2009, p. 57). *De dicto*-internalism, then, is a claim about our concept "moral judgment" and *de re*-internalism a claim about the subject matter—the phenomenon of moral judgments.

Returning to Prinz (and to Roskies's proposal), we can now see that there are really two empirical questions we can ask: First, what is our concept of "moral judgment"? And second, what are moral judgments? Traditionally the first was not regarded as an empirical question. Philosophers probed their intuitions and just assumed that others shared them. Prinz, on the other hand, regards these kinds of questions as empirical in nature and presents his own survey studies that probes *folk intuitions*. He concludes that most people do consider emotions necessary for moral judgments (Prinz this collection, p. 10; for other studies on this with different results see also Nichols 2002, p. 22; Strandberg & Björklund 2013, p. 325; Björnsson et al. 2014, p. 16).

These studies can answer the first question regarding our concept of moral judgments. But, as Prinz rightly points out, people could be wrong (Prinz this collection, p. 10). These studies do not tell us anything about the subject matter. This is a further point Tresan makes. He argues that even if we have internalist intuitions this is not enough to support internalism. He argues that strength in modality is not interesting for a substantial theory of moral opinions. A claim with strong modality doesn't tell us more about the subject of the claim than the same claim without it. That, necessarily, bachelors are unmarried (*de dicto*) tells us nothing more than that they need to be unmarried to be called bachelors. It's a claim about our concept "bachelor". It tells us simply that the subjects are unmarried—the same as this exact claim without modality tells us. But if bachelors were necessarily unmarried (*de re*) this would be bad news for the subjects and would tell us something substantial about them—that they're essentially unmarried, that they, the individuals, are unable to be married. He concludes that "[i]f we are interested in the nature of the Subject Matter, we must look to Relation not Modality" (Tresan 2009, p. 57; emphasis in original).

Only an internalist claim with a strong relation is interesting. But Tresan thinks that there are no arguments for a *de re*-internalism, which would tell us something interesting and substantial about the subject matter. A *de re*-internalism that states a strong Relation is wrong. This is because our intuitions regarding moral judgments and motivation can only support a *de dicto* internalism (Tresan 2009, p. 64). And traditional arguments for internalism provoke only such intuitions.

I think it is clear that Tresan misses one important possible source of evidence for a strong relation: empirical evidence. Here lies the connection to Prinz's work. The empirical findings, which he collected, all point to a strong relation between moral judgment and affective states. I take Prinz to be looking for a strong

Relation when he says that emotions are an "essential component" of moral judgments (Prinz this collection, p. 12).

What I have tried to show here is the following. Prinz raises a potential objection against his own sentimentalism: the relation between moral judgments and emotions lack modal strength. He answers by saying that we have enough evidence to conclude their necessary connection. I argued that this is not a satisfying answer because it misses the core of the objection.

I think the evidence that he has collected points to a strong Relation between moral judgments and affective (motivational) states. Therefore Prinz has an answer to objections that call this strong relation into question. But this is not an answer to an objection that operates with a *de dicto* internalism. Underlying these objections is an opposition to methodological naturalism in general. Antti Kauppinen is one example of someone holding this position (2008, p. 4). Kauppinen does not think we should ask what moral judgments *actually* are. In his view, metaethics is concerned with what moral judgments *necessarily* are. "This takes us from the realm of the actual to the realm of the metaphysical or conceptually possible, and thus beyond the empirical and the observable" (Kauppinen 2008, p. 22).

The evidence that Prinz presents in the target paper doesn't suffice to refute this position. But I hope to have shown that this need not be a cause of concern for Prinz, because this kind of necessity takes us away from the subject matter. At the heart of Prinz's account lies an interest in moral judgments as a natural phenomenon that we should study by empirical means.

Adina Roskies, on the other hand, is sympathetic to empirical philosophy. One of her aims in the internalist–externalist debate was to show that "[...] moral philosophy need not be, and perhaps ought not be, exclusively a priori" (Roskies 2003, p. 2003).

But this is in contrast to her understanding of the required modality of the internalist claim, as I tried to show using Tresan's analysis. If we want to clarify those kinds of debates empirically, it's not enough to just take traditional philosophical claims and look for evidence in their favor or evidence that can refute them. We have to formulate them as a posteriori synthetic claims that are part of a bigger explanatory project (Björnsson 2002, p. 329).

I hope that this can shed more light on the implications of naturalistic metaethics for philosophical claims. They shouldn't be regarded as conceptual a priori claims, but as hypotheses that need empirical confirmation. Naturalistic metaethics is not concerned with a priori conceptual necessities. It requires revising our concepts when they don't fit into the best theories. In that sense empirical philosophers should be revisionists (see Francén 2010, pp. 137 and 142 for a more detailed account of revisionism).

Before I go on, I want to offer one last thought about this. What might be the motivation for framing these positions as claims about conceptual necessity? Roskies writes:

> I take it that internalist philosophers have intended to offer something stronger than contingent claims about human wiring (...) Only a view involving necessity or intrinsicality can distinguish moral beliefs and judgments from other types by their special content. (2008, p. 193)

But why do we need a priori conceptual necessities to distinguish between different kind of beliefs and judgments? We could start with very simple observations. Apparently people play a game of blaming and blessing: they use words like "good" and "bad" that are somehow different than other terms. The task of defining what morality is could be a descriptive anthropological enterprise. And I think this is in the spirit of naturalistic metaethics.

I have argued that it is enough for Prinz's sentimentalism (and for internalism) to claim a strong Relation between moral judgments and emotions. But what kind of Relation is strong enough for it? A mere statistical connection is surely not enough. If the important part of the sentimentalist thesis is not the Modality of the whole claim, we have to analyze the terms "ne-

cessary" and "essential" in a non-modal way. One possibility, that harmonizes with naturalized metaethics, is to regard this connection as *functional*.[5]

In the next, and final, section I shall look at Prinz's critique of non-cognitivism. I shall present a speculative alternative to his view that I hope, again, is in agreement with his proposal for a naturalized metaethics.

3 Defending non-cognitivism as an empirical theory

Here, I want to argue against Prinz's attack on non-cognitivism. He thinks that there are good empirical reasons to reject non-cognitivism. His first argument is that cognitivism can predict the surface form of moral language better than non-cognitivism. First, I argue against this by pointing to non-cognitivist accounts of moral language that I think can predict this surface form. Second, I provide a speculative non-cognitivist theory of why moral language has the surface form we can observe. Again, I think my proposal is in agreement with Prinz's naturalized metaethics. I do think, however, that it challenges him to explore the space of possible accounts. My proposal shows, I hope, that the empirical evidence cannot, at this point, decide this question.

The second question in Prinz's decision tree is whether or not moral judgments are truth-apt. Can they be true or false? Theories that answer yes to this question are cognitivistic, while theories that answer negatively are non-cognitivistic. *Non-cognitivism* is a collective term that can refer to many different theories (Shafer-Landau 2003, p. 17). It consists of two theses (Roojen 2013, section 1.1): the first says that moral utterances do not express propositions; they're not truth apt. This is a semantic thesis about moral language. The second thesis says that moral beliefs are not representational. They do not refer to anything in the world. This is a thesis about the mental state of the moral agent. Here Prinz wants to defend cognitivism by providing empirically-informed reasons to reject non-cognitivism. He defines expressivism in the following way (we can think of Expressivism as one form of the first, semantic thesis of non-cognitivism):

> Expressivism $=_{\mathrm{Df}}$ Moral assertions express mere feelings or non-assertoric attitudes, and do not purport to convey facts. (Prinz this collection, p. 7)

Prinz denies both of the two theses that make up non-cognitivism. He argues that the most obvious empirical prediction of non-cognitivism fails, as he thinks that if non-cognitivism was true we would expect our moral language to have a non-cognitive form (Prinz this collection, p. 16). But this is not the case. It seems that our moral language mostly has declarative form.

If this is correct, and if I don't have reasons to disbelieve it, does it mean that non-cognitivism makes wrong predictions? I don't think this is the case. Much of the work in non-cognitivism is dedicated to explaining this apparent tension. But I don't think that this involves "elaborate logics", as Prinz puts it (this collection, p. 16). Rather, most non-cognitivists provide theories about the nature of moral discourse that show that we should expect the surface grammar to be declarative. I don't think that non-cognitivism has or needs to have these "obvious empirical predictions".

The starting point is to look at the way language is used. It is not the literal meaning of ethical terms that are of interest but their *function* (Björnsson 2002, p. 328). Expressivism entails a pragmatist theory of moral language:

> [T]he pragmatist attempts to describe the function that a word, phrase or concept plays in human life, and once he has satisfied his curiosity there, he does not think that there are any further questions to ask about utterances of that sort. (Smyth 2014, p. 608)

Arguably, such a pragmatist view is easier to naturalize because we have the social sciences,

5 For this proposal see Björnsson & Francén Olinder (2013) and Bedke (2009) and Schulte (2012). They detail the idea that we can think of this relation as *teleo-functionalistic*.

which offer large toolboxes for investigating human practices.

Although Prinz's definition of expressivism may be at the heart of non-cognitivism, in most cases this is not the whole story. According to expressivism, moral terms are not only used to express ones attitudes but also to provoke certain attitudes in the hearer. This idea goes back to the early emotivists. The "dynamic use" of language (Stevenson 1937, p. 21) involves the manipulation of others: "[E]thical terms are instruments used in the complicated interplay and readjustment of human interests" (Stevenson 1937, p. 20; emphasis in original).

Stevenson, and many others following him, analyze expressions like "x is good" as meaning "Hooray for x! Do hooray as well!" (Stevenson 1937, p. 25).[6] It expresses the speakers attitude and the wish or the prescription that the hearer should adopt this attitude as well.

At this point Prinz could reiterate his point and simply ask: "Why then do we say 'this is good' and not 'I like this, do so as well'?" Here I want to offer a speculative answer: because we don't like to be manipulated. If the function of moral language is, at least in part, to influence the attitudes and the behavior of others, I think we should expect it to take this form. This is because a declarative sentence has more *authority* than a mere expressive one. If I want someone to do something it is arguably more effective to disguise it in non-subjective form, to give it the appeal of a truth-aptness.[7] I want to disguise it so that it will serve this persuasive purpose.

I don't want to say that these ideas are correct. But they're plausible theories that predict the surface form of moral language, and which are no worse than cognitivistic theories. Expressivism focuses on what people do with language. It focuses on the speech act, not the literal meaning. Whether people express, declare, prescribe, describe, recommend, or evaluate is nothing we can easily read from the surface form. But this is what Prinz seems to presuppose when he says the most obvious empirical prediction fails. We have to look at their behavior and the pragmatic context in which the discourse happens.

I argue that this fits even better with Prinz's project of a naturalized metaethics. When Prinz discusses the last step in the decision tree, he writes: "Naturalizing relativism will require the marriage of cultural anthropology and sociolinguistics" (this collection, p. 24). I think this marriage could be more helpful at an earlier stage in the decision tree—to help answer the question of whether or not moral terms aim at truth.

4 Conclusion

In this commentary on Prinz's highly interesting and substantial target paper I welcomed his methodological naturalism, but argued that his project is not as conciliatory between traditional analytical philosophy and naturalized philosophy as he seems to think. The reason is that on closer scrutiny we find opposing views on the methodology and purpose of philosophy. In the second part of my contribution I looked at an objection against Prinz's sentimentalism. I argued, first, that he misses the real core of this kind of objections. Then I used Jon Tresan's distinction between *de dicto*- and *de re*-internalism as a conceptual tool to propose and develop another answer that Prinz could use against this objection. In particular, I claimed that, given Prinz's metaethical naturalism, we should not look for conceptual necessity but for fruitful hypotheses which we can test in *a posteriori*. In the third and last part I argued against Prinz's critique of non-cognitivism. Prinz thinks that the most obvious empirical prediction of non-cognitivism fails. Here, I tried to demonstrate how non-cognitivism, given a pragmatical view of moral language, actually predicts the surface grammar of moral discourse as well as cognitivistic alternatives. I proposed a speculative explanation for this interesting fact. This kind of explanation, I believe, fits even better with Prinz's project of a naturalized metaethics.

6 Stevenson (1937, p. 25) writes: "I do like this; do so as well!" But the first part looks suspiciously descriptive. Because this doesn't fit with Stevenson's account, I reformulated it in this way.

7 Mackie discusses this instrumental use when he discusses why people give their moral judgments the appeal of objectivity (1990, p. 42). But as we saw, Prinz thinks this premise is wrong.

References

Bedke, M. S. (2009). Moral judgment purposivism: Saving internalism from amoralism. *Philosophical Studies*, *144* (2), 189-209. 10.1007/s11098-008-9205-5

Björklund, F., Björnsson, G., Eriksson, J., Francén Olinder, R. & Strandberg, C. (2012). Recent work on motivational internalism. *Analysis*, *72* (1), 124-137. 10.1093/analys/anr118

Björnsson, G. & Francén Olinder, R. (2013). "Internalists beware" - We might all be amoralists! *Australasian Journal of Philosophy*, *91* (1), 1-14. 10.1080/00048402.2012.665373

Björnsson, G. (2002). How emotivism survives immoralists, irrationality, and depression. *Southern Journal of Philosophy*, *40* (3), 327-344. 10.1111/j.2041-6962.2002.tb01905.x

Björnsson, G., Eriksson, J., Strandberg, C., Francén Olinder, R. & Björklund, F. (2014). Motivational internalism and folk intuitions. *Philosophical Psychology*, 1-20. 10.1080/09515089.2014.894431

Darwall, S. L. (1995). *The british moralists and the internal 'ought', 1640-1740.* Cambridge, UK: Cambridge University Press.

DeLapp, K. M. (2011). Metaethics. *The Internet Encyclopedia of Philosophy.* http://www.iep.utm.edu/metaethi/

Francén, R. (2010). Moral motivation pluralism. *Journal of Ethics*, *14* (2), 117-148. 10.1007/s10892-010-9074-y

Kauppinen, A. (2008). Moral internalism and the brain. *Social Theory and Practice*, *34* (1), 1-24. 10.5840/soctheorpract20083411

Mackie, J. L. (1990). *Ethics : Inventing right and wrong.* London, UK: Penguin.

Nichols, S. (2002). How psychopaths threaten moral rationalism: Is it irrational to be amoral. *The Monist*, *85* (2), 285-303.

Prinz, J. (2006). The emotional basis of moral judgments. *Philosophical Explorations*, *9* (1), 29-43. 10.1080/13869790500492466

——— (2008). Empirical philosophy and experimental philosophy. In J. Knobe & S. Nichols (Eds.) *Experimental philosophy* (pp. 189-208). Oxford, UK: Oxford University Press.

——— (2015). Naturalizing metaethics. In T. Metzinger & J. M. Windt (Eds.) *Open MIND.* Frankfurt a. M., GER: MIND Group.

Roskies, A. (2003). Are ethical judgments intrinsically motivational? Lessons from "Acquired Sociopathy". *Philosophical Psychology*, *16* (1), 51-66. 10.1080/0951508032000067743

——— (2008). Internalism and the evidence from pathology. In W. Sinnott-Armstrong (Ed.) *Moral psychology: The neuroscience of morality: emotion, brain disorders, and development* (pp. 191-206). Cambridge, MA: MIT Press.

Schulte, P. (2012). Satan Und Der Masochist: Eine Nonkognitivistische Antwort auf den Amoralismus-Einwand. In A. Dunshirn, E. Nemeth & G. Unterthurner (Eds.) *Crossing Borders. Grenzen (über)Denken. Beiträge Zum 9. Internationalen Kongress der österreichischen Gesellschaft für Philosophie in Wien* (pp. 599-608). Wien, AUT: Österreichische Gesellschaft Für Philosophie.

Shafer-Landau, R. (2003). *Moral realism: A defence.* Oxford, UK: Oxford University Press.

Smyth, N. (2014). Resolute expressivism. *Ethical Theory and Moral Practice*, *17* (4), 607-618. 10.1007/s10677-014-9495-y

Stevenson, C. L. (1937). The emotive meaning of ethical terms. *Mind, New Series*, *46* (181), 14-31.

Strandberg, C. & Björklund, F. (2013). Is moral internalism supported by folk intuitions? *Philosophical Psychology*, *26* (3), 319-335. 10.1080/09515089.2012.667622

Tresan, J. (2009). Metaethical internalism: Another neglected distinction. *Journal of Ethics*, *13* (1), 51-72.

van Roojen, M. (2013). Moral cognitivism vs. non-cognitivism. *The Stanford Encyclopedia of Philosophy*, *Winter 2013* E. N. Zalta (Ed.) http://plato.stanford.edu/archives/win2013/entries/moral-cognitivism/