Should Metaethical Naturalists Abandon *de dicto* Internalism and Cognitivism?

A Reply to Yann Wilhelm

Jesse Prinz

Yann Wilhelm pursues three issues in response to my target article. First, he tries to expose my naturalism as more radical than I let on. I concede the point, though I also offer ways in which my radicalism might be mitigated. Second, he exposes a limitation in my argument for internalism, and suggests that naturalists should defend form on internalism that is neutral about conceptual claims (*de re* internalism, rather than *de dicto*). I welcome the suggestion, but also consider how naturalists might defend *de dicto* internalism. Third, Wilhelm challenges my argument against non-cognitivism, by offering a novel explanation of the fact that moral judgments have an assertoric form. I response, I note avenues for cognitivist resistance to Wilhelm's explanation.

Author

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

Commentator

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

Editors

Thomas Metzinger
metzinger@uni-mainz.de
Johannes Gutenberg-Universität
Mainz, Germany

Jennifer M. Windt
jennifer.windt@monash.edu
Monash University
Melbourne, Australia

1 Introduction

In "Naturalizing Metaethics," I try to establish that core questions in metaethics lend themselves to empirical investigation. I argue that we can potentially adjudicate long-standing debates by testing predictions made by competing metaethical theories. I also make some conjectures about how such empirical investigations will turn out. Based on a small selection of preliminary findings, I advance a case of a version of sentimentalism—the view that emotions are essential to moral judgments. I also suggest that sentimentalism commits me to internalism—the view that moral judgments are essentially motivating—and I advance an empirical case for cognitvism—the claim that moral judgments are, like other assertions, capable of being true or false.

In his insightful commentary, Yann Wilhelm offers clarifications and challenges to my

arguments. First, he asks whether my naturalism is compatible with traditional approaches in philosophy. I imply that the two can co-exist in a complementary way, but Wilhelm suggests that my naturalism is more radical that it appears. I am forced to agree, and to clarify the co-existence claim. Wilhelm also challenges my case for internalism, distinguishing two different forms and suggesting I am only in a position to argue for one of them. I am open to that possibility, but I also sketch a strategy for defending both forms. Wilhelm concludes with a challenge to my defense of cognitivism. He provides non-cognitivists with an explanation for findings that I say they cannot explain. I offer a cognitivist response, but grant that this proposal demands empirical attention.

Wilhelm's commentary provides a valuable contribution to empirically oriented metaethics. He offers strategies for avoiding certain kinds of debates with opponents of naturalism, and he identities empirical issues that can be used settle debates between card-carrying naturalists. Wilhelm deepens my understanding of these issues and strengthens my optimism about the prospects of naturalistic metaethics.

2 Is naturalism a radical position?

Before moving on to the first order debates that Wilhelm so helpfully pursues, I want to concede an important point that he makes in the opening of his commentary. Wilhelm rightly observes that I overstate the extent to which a thoroughgoing naturalism can preserve traditional approaches to philosophy. Though ostensibly a plea for consilience, I am, in fact, skeptical about the notion *a prioricity*. Rather, I claim that armchair methods are observational (intuitions are defeasible inner observations informed by prior experience, and open to empirical correction). As Wilhelm makes clear, traditionalists who view conceptual analysis as an *a priori* endeavor will not share my enthusiasm for naturalism.

In another respect, however, my position is conservative. I don't think traditional philosophers must stop working as they currently do. Armchair methods remain the primary source of

philosophical theories and distinctions. They also are the primary source for philosophical thought experiments that can be used to test between theories. Thus, my invitation to interpret armchair methods as observation is intended as a vindication of traditional philosophy, though not a vindication of how some traditional philosophers understand their own endeavors.

Proof of this qualified vindication comes form the fact that empirically oriented philosophers regularly draw on traditional work in devising their studies. For example, experimental philosophers have used trolley problems, twin earth cases, and the thought experiments used to back contextualism in epistemology. In my target paper, I relied on theories that have been identified and articulated within traditional philosophy. Testing between theories requires observation, I believe, but it would be a great loss if every philosopher ran a laboratory. Instead, I envision a future for philosophy in which many researchers do no experimental work, others are primarily experimentalists, and still others do a combination of the two. If we begin to make empirical methods a standard part of philosophical training, then philosophers will be able to read psychological research more responsibly and conduct experiments when they see fit. But it doesn't mean that they will also suddenly stop thinking and blindly collect data. As in the sciences, theoretical work in required in philosophy. We can resist the idea of *a priori* truth without throwing away the armchair.

3 Must naturalist be content with *de re* internalism?

These methodological points bear on Wilhelm's first challenge to my metaethical conclusions. In the target paper I argue for a form of internalism (roughly, the view that moral judgments are essentially motivating). Wilhelm points out that my evidence for this claim will not satisfy many externalists. I primarily rely on evidence that moral judgments always co-occur with emotional states, but fee externalists will be impressed; they will say that such findings cannot address questions about whether it is necessar-

ily the case that moral judgments are motivating, even if they always happen to be are motivating.

Wilhelm helpfully replies to this objection on my behalf, using Jon Tresan's distinction between *de dicto* and *de re* internalism. The former is a thesis about the concept of moral judgment (viz., it is a conceptual truth that when that concept applies, motivation applies as well). The latter is a claim about moral judgments themselves (viz., moral judgments do in fact carry motivation force). Wilhelm concurs that my evidence can contribute to a defense of the *de re* claim. He suggests that I abandon the case for *de dicto* internalism, since naturalists should not concern themselves with conceptual claims.

I welcome Wilhelm's suggestion, and I am inclined to endorse it. Let me mention, however, a strategy available to the naturalist whose heart is set on defending the *de dicto* claim. Returning to Wilhelm's discussion of methodology, let's imagine that naturalists wage a successful campaign against the *a priori*. Properly pursued, such a campaign might also undermine metaphysical necessity. Metaphysical necessities, unlike nomological necessities, are alleged to be true in virtue of conceptual entailments rather than laws of nature or natural facts. The critique of *a prioricity* threatens metaphysical necessity because it advances the view that truths about concepts are open to empirical revision. Let's suppose that concepts are mental representations garnered through experience with the function of classifying things in the world. So construed, concepts are susceptible to improvement through empirical inquiry. Initial concepts are rough and ready pointers that we use to carve up the observational world, and revised concepts are carvings that remain after observation. Now let us define a "robust conceptual truth" as the conceptual entailments that survive after a concept has been subjected to empirical fine-tuning. Such truths would more or less coincide with how the world is, together with certain pragmatic assumptions that go into theory construction. Thus, they would coincide with truths that emerge from our study of the things themselves (which are also constrained by pragmatic assumptions). On this picture, *de dicto* collapses into *de re*. A defense of *de re* internalism would indicate that our concept of moral judgment will converge on internalism as well. Rather than bypassing *de re* internalism, we can try to defend it by naturalizing conceptual truth.

Wilhelm might reply that this defense of *de dicto* internalism would not persuade non-naturalists. The defense is based on the assumption that the naturalist critique *of a prioricity* goes through, but that is just what non-naturalists are inclined to deny. Thus, it might appear the that the debate over the *de dicto* position is hostage to unresolvable disputes about the nature of philosophy.

Here I'd balk at the claim that such disputes are unresolvable. Those who believe is *a prioricity* may dislike naturalism, but they certainly believe that their views require evidential support. Naturalists offer an account of what concepts are (mental representations) and an explanation of conceptual intuitions (introspection of mental representations). Non-naturalists are obliged to provide an alternative account of both, and the two accounts can then be compared by agreed upon standards. I venture that the naturalist account will find a resounding victory in such a head-to-head match. It is more parsimonious view, since both sides must grant the existence of mental representations, and I suspect it can fully account for our conceptual intuitions.

These are, of course, big debates, which I cannot settle here. My point is simply that we can imagine a two-stage process that begins with broad issues about naturalism, and then moves on to first-order views. On my prognosis, we won't end up abandoning the notion of conceptual truth, but rather revising it. If so, *de dicto* naturalism might turn out true. Wilhelm may be right, however, that until we come to greater consensus on the nature of philosophy, naturalists might be on firmer ground if they try to bypass conceptual questions. He is also right that, from a naturalist perspective *de re* internalism may be the more interesting thesis. Conceptual claims lose their distinctive interest if concepts are revisable and, ultimately, coincident with empirical theories.

4 Can non–cognitivists explain the assertoric form of moral judgments?

Let me turn, finally, to Wilhelm's constructive effort to defend non-cognitivism. Non-cognitivists claim that moral judgments are not like ordinary assertions; they cannot be assessed as true or false, but rather merely express the speakers attitudes and commendations. If so, I ask, why do we express moral judgments as assertions? This is a familiar challenge. In my discussion, I merely I point out that can be backed up by empirical data. Wilhelm has a two-part reply. First, he observes that, for non-cognitivists, the primary function of moral discourse is to persuade. C. L. Stevenson, for example, says that "x is bad" does not just mean "boo to x!"; it also means and "say boo to x as well!". Second, Wilhelm makes the original and plausible suggestion that this persuasive function is most effective when it covert. People, he notes, don't like to be manipulated. If I explicitly exhort you to say "boo!" you may resist, because no one likes being told what to do. But if I present my attitude in the form on an assertion, you might causally take it on board, as you would if I were presenting an ordinary statement of fact.

I think Wilhelm's proposal deserves serious exploration. Cognitivists can respond in two ways. First, they can try to show that moral discourse often occurs in contexts that don't aim at persuasion. This might seem implausible. After all, why should we bother engaging in moral discourse if we don't intend to persuade anyone? On closer analysis, however, it does seem that much of our moral discourse involves preaching to the choir. In political debates, for example, left wing pundits and right wing pundits engage in a lot of moral discourse, but they never seem to persuade each other. This raises the intriguing possibility that moral judgments are not primarily in the business of persuasion. An alternative possibility is that we make moral judgments to assert our identity, or express solidarity with like-minded individuals. Empirical tests might be designed to compare the persuasion model and the self-expression model.

Cognitivists might also try to resist Wilhelm's conjecture that people do not like to be manipulated by consulting research on explicit persuasion. In defense of Wilhelm's conjecture, there is a literature suggesting that people sometimes resist explicit persuasion (e.g., Petty & Cacioppo 1979). On the other hand, resistance does not occur in all contexts. Indeed, in a consumer product context, Reinhard et al. (2006) found that, when a person is regarded as likeable (or attractive!), they become even more persuasive when they make their intent to persuade explicit. Similarly, in studies of college drinking behavior, Neighbors et al. (2008) found that injunctive norms (which explicitly reference attitudes) are effective when and only when they are expressed by members of the students' social groups. Further work could test the effects of explicit injunctions in the moral domain.

I should underscore that I think more testing is required to settle these debates. Wilhem's explanation for surface discourse remains viable, and we can make progress on these issues by devising new ways to test it. These are manifestly empirical issues. While I wager with the cognitivists, I grant that the case is far from closed.

5 Conclusion

I am indebted to Yann Wilhelm for his generous and probing commentary. It brings welcome clarification and new challenges to the project I set out "Naturalizing Metaethics." I also welcome the spirit of Wilhelm's discussion, which moves beyond ideological debates about metaphilosophy, and offers promising strategies for answering core metaethical questions.

Wilhelm successfully establishes that my preferred form of naturalism is less compatible with traditional philosophy than I let on, but I also pointed out that work by traditionally minded philosophy remains an invaluable font of philosophical theories. Wilhelm then offers a helpful suggestion that naturalists might more easily defend internalism if they bypass conceptual versions of that view. In response, I suggested that the radical implications of naturalism may actually offer a way to defend the concep-

tual version of internalism, by advancing a naturalized account of conceptual truth. Finally, Wilhelm offered a new psychological cum functional account of moral discourse, which inoculates non-cognitivists against grammatical objections. While I hold out hope for cognitivism, Wilhelm has identified a genuine empirical challenge to the cognitivist. This challenge beautifully demonstrates the value of empirical testing in metaethics, and it also reminds us that there is much work to be done.

References

Neighbors, C., O'Connor, R. M., Lewis, M. A., Chawla, N., Lee, C. M. & Fossos, N. (2008). The relative impact of injunctive norms on college student drinking: The role of reference group. *Psychology of Addictive Behaviors: Journal of the Society of Psychologists in Addictive Behaviors*, *22* (4), 576-581. 10.1037/a0013043

Petty, R. E. & Cacioppo, J. T. (1979). Effects of forewarning of persuasive intent and involvement on cognitive responses and persuasion. *Personality and Social Psychology Bulletin*, *5* (2), 173-176. 10.1177/ 014616727900500209

Reinhard, M.-A., Messner, M. & Sporer, S. (2006). Explicit persuasive intent and its impact on success at persuasion – The determining roles of attractiveness and likeableness. *Journal of Consumer Psychology*, *16*, 249-259. 10.1207/s15327663jcp1603_7