



DISCUSSION NOTE

INTERNALISM, IDEAL ADVISORS AND THE CONDITIONAL FALLACY

BY ALEXANDER HYUN

JOURNAL OF ETHICS & SOCIAL PHILOSOPHY

DISCUSSION NOTE | AUGUST 2015

URL: WWW.JESP.ORG

COPYRIGHT © ALEXANDER HYUN 2015

Internalism, Ideal Advisors and the Conditional Fallacy

Alexander Hyun

IN HER RECENT ARTICLE, “INTERNALISM ABOUT REASONS: Sad But True?” Kate Manne offers a brilliant defense of a novel version of internalism about normative reasons. But I will argue that this defense is not successful. After explaining the nature of Manne’s internalism, I offer two counterexamples to it, thereby showing that her argument in its favor goes wrong somewhere. I then identify the false premise in her argument. In brief, I suggest that Manne’s “practice-based approach” to practical normativity (or at least, to normative reasons) should be rejected.

1. Manne’s Internalism About Reasons

Manne’s internalism about reasons is as follows:

Manne’s Internalism (MI): An agent A has a reason to Φ only if A would end up in a state of being (somewhat) motivated to Φ , following an idealized process of being reasoned with about the matter of what to do.¹

Manne understands the “idealized process” mentioned in this principle to be a completed conversational process in which A ’s “ideal advisor” reasons with A about what A ought to do.² We are supposed to imagine A ’s ideal advisor as being “a flesh and blood human being” who is “possessed of all the relevant information” and who is “fully procedurally rational (or at least as fully procedurally rational as any actual human being could be).”³ Further, A ’s ideal advisor is virtuous, wise, well-disposed toward her advisee and ideally suited for “getting through” to A , morally.⁴ According to Manne, then, an idealized process of an agent’s being reasoned with about the matter of what to do is a completed process in which an advisor with all of these features reasons with the agent about what she ought to do.

2. Counterexamples to Manne’s Internalism

A counterexample to MI is a case in which an agent A has a reason to Φ , yet it is false that A would end up in a state of being (somewhat) motivated to Φ , following an idealized process of being reasoned with about the matter of what to do. Consider two cases of this sort.

¹ Manne (2014: 109).

² Ibid.: 97.

³ Ibid.

⁴ Ibid.

2.1. Extreme introverts

Suppose that Ms. Introvert is just like any ordinary human being in this respect: If she goes too long without having a conversation with another person, she starts to feel terribly lonely and sad. But further, suppose that Ms. Introvert is unusually introverted, in this way: All social interaction thoroughly exhausts her, such that it takes weeks for her to psychologically recuperate after conversing with someone. In fact, when she has more than one conversation in a month, she becomes completely miserable.

Now, suppose that Ms. Introvert has not interacted with anybody for a very long time, and she starts to feel lonely. Fortunately, her best friend from high school is in town, and she knows that a visit with her old friend would buoy her spirits. Intuitively, Ms. Introvert has an excellent reason to call her friend and arrange to meet up with her. Yet it is surely false that Ms. Introvert would end up in a state of being motivated to call her friend, following an idealized process of being reasoned with over the matter of what to do. For, following such a process, Ms. Introvert would have had her fill of socializing, and calling her old friend would be the last thing she would want to do. So, Ms. Introvert's case is a counterexample to MI.

2.2 Extremely proud agents

Mrs. Prideful prides herself on her rationality, and she cannot stand the thought that there might be people who are more rational than she is. In fact, she has the following extreme disposition: Whenever it becomes clear to her that she is interacting with someone who is far more procedurally rational than she is, she becomes overwhelmingly depressed, losing all of her ordinary cares and concerns. Fortunately, there are no actual human beings who are far more rational than Mrs. Prideful, so she is able to lead a completely ordinary life. Mrs. Prideful is bizarre, but it certainly seems possible for there to be an agent with her psychological profile.

Now, suppose that it is an ordinary day, and Mrs. Prideful starts to feel famished. Intuitively, she has a reason to eat a snack. Yet it is surely false that Mrs. Prideful would end up in a state of being motivated to get a snack, following an idealized process of being reasoned with over the matter of what to do. For, following such a process, Mrs. Prideful would realize that she is interacting with someone who is far more procedurally rational than she is, and she would therefore have none of her ordinary motivations, such as the motivation to eat when she is hungry. So, Mrs. Prideful's case is another counterexample to MI.

2.3 The conditional fallacy

Another way to put the objections to MI that I have offered is that MI commits the conditional fallacy, a well-known problem for analyses of rea-

sons according to which the possession of a reason depends on the truth of some counterfactual. As Robert Johnson helpfully explains:

The fallacy consists of overlooking, in various ways, dependencies between the *analysandum* and the antecedent and consequent of the conditional in the *analysans*. For instance, one might overlook a set of counter-examples for one's analysis in which the *analysandum* is obviously true, yet because the antecedent of the conditional in the *analysans* is in some way incompatible with the consequent in the *analysans*, the *analysans* is false.⁵

The above two counterexamples show that MI commits the conditional fallacy in this way. Consider again the case of Ms. Introvert. It is obviously true that Ms. Introvert has a reason to call her friend; yet, because Ms. Introvert's conversing with an ideal advisor is incompatible with her subsequently having a motivation to call her friend, the following counterfactual is false: If Ms. Introvert were reasoned with by an ideal advisor about what to do, then she would end up motivated to call her friend. Similarly, it is obviously true that Mrs. Prideful has a reason to eat a snack; yet, because her conversing with an ideal advisor is incompatible with her subsequently having a motivation to eat a snack, the following counterfactual is false: If Mrs. Prideful were reasoned with by an ideal advisor about what to do, then she would end up motivated to eat a snack. By overlooking the ways in which the antecedent and consequent of MI's consequent can conflict, Manne overlooks these counterexamples.

And now that we see that I have basically been relying on the conditional fallacy to advance my criticisms, it may be thought that Manne has replies. In a lengthy footnote, Manne offers some thoughts on her general strategy for addressing conditional-fallacy worries.⁶ She suggests that we may "prevent some of the well-known conditional fallacy worries which afflict 'ideal agent' models from afflicting [her] account too" by being "very careful about how we individuate conversations" that transpire between an agent and her ideal advisor. To illustrate, suppose that one raises the following conditional-fallacy worry for Manne's view. Mr. Thirsty wants to drink gin and he believes that the glass in his hand is filled with gin. But, in fact, the glass is full of petrol, which Mr. Thirsty would hate to drink. Intuitively, Mr. Thirsty has a reason to inquire into the contents of the glass. But the following counterfactual seems false: If Mr. Thirsty were reasoned with by an ideal advisor, then he would be motivated to inquire into the contents of the glass. For at the end of a conversation with his fully informed advisor, Mr. Thirsty will have been told about the true contents of the glass, and thus will not be the least bit motivated to inquire into the matter.

⁵ Johnson (1999: 54).

⁶ Manne (2014: 105).

Manne replies to this conditional-fallacy worry by suggesting that “conversations should be individuated in a more fine-grained way.”⁷ There is a (short) conversation between Mr. Thirsty and his ideal advisor that goes something like this:

Ideal Advisor: The stuff in that glass is not what you think it is. Would you like to know what’s really in it?

Mr. Thirsty: Yes.

Ideal Advisor: Then I recommend that you inquire into the contents of the glass.

Since Mr. Thirsty is motivated to inquire into the contents of the glass by the end of this short conversation between him and his ideal advisor, Manne suggests that her view does not entail the counterintuitive claim that Mr. Thirsty does not have a reason to inquire into the contents of the glass. It therefore seems that Manne’s considered judgment is that an agent has a reason to Φ only if there is *some* (perhaps short) conversational process of being reasoned with by an ideal advisor about what to do, such that if the agent were to go through this process, then she would end up in a state of being motivated to Φ . And on this refined statement of MI, the above conditional-fallacy worry is not a problem.

I agree that this general strategy for responding to conditional-fallacy worries provides a plausible line of response to the specific conditional-fallacy worry just discussed. But it seems hopeless as a response to the conditional-fallacy worries that I offered above. For Ms. Introvert (we may now stipulate) becomes exhausted after even the shortest of social interactions, and Mrs. Prideful becomes depressed immediately upon conversing with somebody who is superior to her with respect to procedural rationality. So, Manne cannot avoid the specific conditional-fallacy worries that I have raised by being careful about how she individuates conversations between agents and their ideal advisors.

3. Rejecting Manne’s Practice-Based Approach to Practical Normativity

I have argued that MI is false. Since Manne gives an intriguing argument for her view, it is incumbent on me to identify where I take it to go wrong. Her argument is as follows:

- (1) A reason for an agent A to Φ is a consideration that would be apt to be cited in favor of A’s Φ -ing by her ideal advisor, who is reasoning with her in an ideal way about what she ought to do.
- (2) There is a consideration that would be apt to be cited in favor of A’s Φ -ing by her ideal advisor, who is reasoning with her in an ideal way about what she

⁷ Ibid.: 105.

ought to do, only if A would end up in a state of being (somewhat) motivated to Φ following an idealized process of being reasoned with in this way.

- (3) So, an agent A has a reason to Φ only if A would end up in a state such that she would be (somewhat) motivated to Φ , following an idealized process of being reasoned with in this way.⁸

My position is that premise (1) is false. Underlying premise (1) is Manne's "practice-based approach" to practical normativity. On this approach, abstract normative notions are analyzed in terms of their role within some type of *normative behavior* – that is, "behavior by means of which we give voice to ideas about what to do, and also what should happen."⁹ Premise (1) expresses the thought that the concept of a *reason* is correctly analyzed in terms of its role within the normative behavior of reasoning with someone about what to do. But this premise should be rejected.

First, consider Manne's support for premise (1). So far as I can see, she gives only the following argument in favor of it:

The above proposal [that reasons just are the considerations that would be apt to be cited by one's ideal advisor] seems to me attractive partly insofar as it secures a close connection between reasons for action and the activity of reasoning with a person about what she ought to do. It is hard to believe that the entities and the activity could come too far apart. Surely the connection goes deeper than the common etymological root of the corresponding English words.¹⁰

I fully agree that there must be some non-etymological explanation of the fact that there is a "close connection" between reasons and the activity of reasoning with a person about what she ought to do. I also agree that premise (1) provides such an explanation. But this counts as a strong reason to accept (1) only if it is the best explanation of the *explanandum*, and it is far from clear that this is so. While Manne proposes to explain this connection by analyzing reasons in terms of the activity of reasoning, we could alternatively explain this connection by analyzing the activity of reasoning in terms of reasons. For instance, the following account of the nature of reasoning with someone seems plausible:

Reasoning: Agent A_1 's reasoning with agent A_2 about what A_2 ought to do just is the activity in which A_1 helps A_2 come to recognize some consideration as a reason for action.

This analysis accounts for the fact that reasons and reasoning with someone are closely connected, and it obviously does not require the truth of premise (1). The availability of this account therefore greatly reduces the strength of Manne's argument for premise (1).

⁸ Manne (2014: 109). Manne presents her argument in more detail, but the details do not matter for my purposes.

⁹ Ibid.: 94.

¹⁰ Ibid.: 98.

In addition to the failure of Manne's support for premise (1), there are decisive reasons to reject it. Premise (1) entails the following:

- (A) If R is a reason for an agent A to Φ , then R is a consideration that would be apt to be cited in favor of A's Φ -ing, by her ideal advisor, who is reasoning with her in an ideal way about what she ought to do.

And (A) is subject to the same sorts of counterexamples that afflict MI. To illustrate, consider for a final time the case of Ms. Introvert. The following is surely true: *That Ms. Introvert is lonely and craving social interaction* is a reason for her to call her friend. But the following is false: *That Ms. Introvert is lonely and craving social interaction* is a consideration that would be apt to be cited in favor of Ms. Introvert calling her friend, by her ideal advisor, who is reasoning with Ms. Introvert in an ideal way about what she ought to do. For, once the ideal advisor is reasoning with Ms. Introvert, Ms. Introvert is neither lonely nor craving social interaction, and it would not be apt for an ideal advisor to appeal to a false claim while reasoning with Ms. Introvert about what to do. (A) is therefore false. And so premise (1) should be rejected.

4. Conclusion

Manne holds that her internalism implies that the callous husband who beats his wife may have no reason to change his ways.¹¹ This is why she confesses that her internalism is sad, despite being likely true.¹² But I conclude that we have reason to be glad: Happily, Manne's Internalism about reasons is false, and we can go on believing that the callous husband has every reason to treat his wife more kindly.¹³

Alexander Hyun
University of Wisconsin–Madison
Department of Philosophy
abhyun@wisc.edu

¹¹ Ibid.: 111.

¹² Ibid.: 92.

¹³ I am grateful to the following people for helpful feedback on this paper: Megan Hyun, Eric Sampson, Ben Schwan, Russ Shafer-Landau and an anonymous reviewer of this journal.

References

- Johnson, R. (1999) "Internal Reasons and the Conditional Fallacy," *Philosophical Quarterly* 49: 53-71.
- Manne, K. (2014) "Internalism About Reasons: Sad But True?" *Philosophical Studies* 167: 89-117.