DIRECTION OF FIT AND THE AIM OF DESIRE

ABSTRACT: Does the desire to perform an action represent that action as good? This question is important because many action-theorists rely on the claim that a desire represents its object as good in order to argue that a desire to perform an action may justify that action. This approach is typically adopted by advocates of the Guise of the Good theory, the thesis that desire aims at the good. According to most Guise of the Good theorists, a desire aims at the good because a desire represents its object as good. In this paper, I show that the claim that a desire represents its object as good entails that the desire has mind-to-world fit. This is problematic since most Guise of the Good theorists who claim that desires represent their objects as good are also committed to saying that a desire aims at the good in virtue of having world-to-mind fit. Contra such theorists, I argue that if we assume that the aim of a desire is connected to the fact that it has world-to-mind fit, then it is plausible to hold that the aim of desire implicates not the good, but the attainable. In the process of defending my conclusion, I offer an alternative interpretation of the direction-of-fit metaphor to that offered by Anscombe.

> The conceptual connexion between 'wanting'…and 'good' can be compared to the conceptual connexion between 'judgment' and 'truth'. Truth is the object of judgment, and good the object of wanting; it does not follow from this either that everything judged must be true, or that everything wanted must be good.
>
> -Anscombe[1]

1. Introduction

Let us refer to the desire to perform an action as an *action-desire*. The Guise of the Good Theory of action-desires (GG theory) is the thesis that action-desires aim at the good.[2] According to the version of GG theory that derives its primary inspiration from the

[1] Anscombe [1957: 76].

[2] Philosophers who have held a view recognisably similar to GG theory include Plato, Aristotle, Aquinas, Kant, Anscombe and Davidson, along with numerous contemporary thinkers. Most contemporary GG theorists hold that not only action-desires, but desires in general, aim at the good. However, insofar as action-desires are a subset of desires, such theorists may also be said to be committed to the claim that action-desires aim at the good. Moreover, by restricting my focus to action-desires, I do not mean to deny that desires in general may be said to aim at the good. At present, I wish to remain non-committal on this question. Furthermore, since most discussions of desire in the GG theory literature (particularly those versions of GG theory that draw their primary inspiration from Anscombe) focus on desires that may potentially lead to or feature in intentional action, my decision to restrict my focus to action-desires is not unprecedented. The main motivation for such a restriction is the recognition that the desire to perform a particular action plays a very different role in our deliberation about what to do than desires that do not have actions as their object. As such, action-desires are deserving of special consideration.

work of Elizabeth Anscombe, the claim that an action-desire aims at the good is taken to entail the following set of claims:

I. An action-desire aims at the good because an action-desire represents the desired action as good.

II. An action-desire is correct only if the desired action is good because an action-desire aims at the good.

III. An action-desire has world-to-mind fit because the world is incorrect or subject to revision when there is a mismatch between the world and what is desired, and a belief has mind-to-world fit because the belief is incorrect or subject to revision when there is a mismatch between what is believed and the world.

IV. The aim of an action-desire explains, is explained by, or is otherwise linked to the fact that it has world-to-mind fit, and the aim of a belief explains, is explained by, or is otherwise linked to the fact that it has mind-to-world fit.

I will refer to the conjunction of the four preceding claims as Anscombean GG theory. Much of the contemporary interest in Anscombean GG theory stems from the fact that it provides us with a framework for making sense of how an action-desire may justify the desired action. According to (I), an action-desire aims at the good in virtue of the fact that it represents the desired action as good.[3] Insofar as representing an action as good may justify performing that action, Anscombean GG theory entails that an action-desire may justify performing the desired action.

In this paper, I will be advancing an alternative to Anscombean GG theory. My aim is both critical and constructive. On the critical side of things, I argue that (III) is an unsatisfactory way of understanding the direction of fit metaphor. Moreover, I argue that, while perhaps not strictly speaking an inconsistency, there is a tension between (I) and (IV). On the constructive side of things, I advance an alternative interpretation of the direction of fit metaphor to that offered by Anscombe, which I refer to as the *two-content interpretation*. If the two-content interpretation is right, then insofar as the aim of an action-desire is linked to the fact that it has world-to-mind fit, it would be more plausible to hold that an action-desire aims at what is attainable than that it aims at the good.[4]

---

[3] There are two different accounts of why desires entail representations of the good currently found in the literature: (i) because desires involve beliefs or judgements of the good (e.g., Price [1989]; Byrne and Hájek [1997]), or (ii) because desires involve perceptions of the good (e.g., Stampe [1987]; Tenenbaum [2007]).

[4] Velleman [1992] also defends a version of the claim that desires aim at the attainable rather than the good. However, the present investigation introduces an important qualification. I hold that it is plausible that action-desires aim at the attainable if we assume that the aim of an action-desire is somehow linked to its direction of fit. Moreover, Velleman's arguments presuppose a conception of direction of fit very similar to that of Anscombe, one that is subject to objections similar to those presented below. This, I believe, significantly weakens Velleman's ability to offer a compelling defense of the claim that desire aims at the attainable.

2. The Anscombe Interpretation

In §32 of *Intention*, Anscombe employs an example comparing a shopping list and a detective's record to illustrate two different ways in which our words (written or spoken) may relate to the world. She writes:

> Let us consider a man going round a town with a shopping list in hand. Now it is clear that the relation of this list to things he actually buys is one and the same whether his wife gave him the list or it is his own list; and that there is a different relation when a list is made by a detective following him about. If he made the list itself, it was an expression of an intention; if his wife gave it to him, it has the role of an order. What then is the identical relation to what happens, in the order and the intention, which is not shared by the record? It is precisely this: if the list and the things that the man actually buys do not agree, and if this and this alone constitutes a mistake, then the mistake is not in the list but in the man's performance...; whereas if the detective's record and what the man actually buys do not agree, then the mistake is in the record.[5]

Let us refer to the example discussed here as the *shopping list example*. Anscombe notes that, in the case of the shopping list, the task is to get the world (the items purchased) to match our words (the shopping list), while in the case of the detective's record, the task is to get the word (the detective's record) to match the world (the items purchased). In the terminology of speech act theorists, the former is an example of *world-to-word* fit, and the latter of *word-to-world* fit.[6]

Significantly, Anscombe sees an intention (i.e., a certain type of psychological attitude) and an order (i.e., a certain type of speech act) as standing in the same logical relation to the man's actions. By her lights, it does not matter if we see the shopping list as an expression of an intention (on the part of the man) or as an order (on the part of his wife); in either case, the shopping list displays the same direction of fit. Thus, Anscombe takes the observation that our words may relate to the world in two very different ways as also applicable to psychological attitudes. This idea has been taken up by several theorists in moral psychology and the philosophy of mind, where a distinction is often drawn between attitudes that display *mind-to-world* fit and attitudes that display *world-to-mind* fit. Huw Price summarises the central intuition behind the direction of fit analysis of psychological attitudes as follows:

> Beliefs have a 'mind-to-world' direction of fit; they aim at fitting the world, at being true. Desires, on the other hand have a 'world to mind' direction of fit; we aim to change the world to fit action-desires, and not vice versa.[7]

---

[5] Anscombe [2000: 56].

[6] The expression "direction of fit" seems to have been first used John Searle, who employed it to describe the two different ways words may relate to the world, as highlighted by Anscombe (See: Searle [1985]).

[7] Price [1989: 120].

The standard way of unpacking the direction of fit metaphor, which builds on Anscombe's shopping list example, is in terms of whether a psychological attitude or the world is deemed to be incorrect or subject to revision when there is a lack of correspondence between the two. According to this approach, what it means for a belief to have mind-to-world fit is that the belief is deemed to be incorrect or subject to revision when there is a mismatch between the world and what is believed. By contrast, what it means for an action-desire to have world-to-mind fit is that the world is deemed to be incorrect or subject to revision when there is a mismatch between that which is desired and the world. Call this the *Anscombe interpretation*.

Unfortunately, the Anscombe interpretation seems to conflate the claim that an attitude has a certain direction of fit with the claim that an attitude has certain correctness conditions. There is widespread agreement that cognitive attitudes like hypothesising that *p*, supposing that *p,* and pretending that *p* have the same direction of fit as believing that *p*. All four attitudes represent things as being a certain way. However, while believing that *p* and hypothesising that *p* are plausibly thought of as incorrect or subject to revision just in case it is not true that *p*, the same cannot be said of the attitudes of supposing and pretending that *p*. Although supposing that *p* and pretending that *p* represent things as being a certain way, neither is necessarily incorrect or subject to revision if *p* is false. Consequently, if we assume, in keeping with the Anscombe interpretation, that an attitude has mind-to-world fit only if it is incorrect or subject to revision when there is a mismatch between the attitude and the world, then we seem forced to deny that supposing and pretending have mind-to-world fit. In order to avoid this difficulty, there may be a temptation to weaken the notion of "incorrectness" implicated by the Anscombe interpretation so that it can plausibly be applied to the attitudes of supposing and pretending. However, there does not seem to be any plausible sense in which pretending is incorrect if things turn out to be different from how they are pretended to be. On the contrary, it would be more plausible to say that pretending indeed aims to represent things in a way that they are not.

The preceding observation suggests a possible diagnosis of where the Anscombe interpretation goes wrong. To say that a particular attitude has mind-to-world fit is not yet to say anything about when that attitude is incorrect or subject to revision. Whether or not an attitude is incorrect or subject to revision seems tied to its aim, and not to its direction of fit. Hence, a belief is incorrect just in case the proposition believed is false because belief aims at truth.[8] Since the attitude of pretending, by contrast, does not aim at the truth, it is not correct just in case the proposition pretended is true. In order to avoid this difficulty, we need to make sense of the direction of fit metaphor in a way that preserves the distinction between saying that an attitude has a certain formal aim (and consequently, certain correctness conditions) and the claim that an attitude has a particular direction of fit. I submit that the Anscombe interpretation is unsatisfactory because it falls short in precisely this respect.

[8] For a discussion of some of the different ways we can understand the claim that belief aims at truth, see Engel [2004].

Putting aside the above difficulty with the Anscombe interpretation, there is an additional problem with Anscombean GG theory that I wish to highlight. When the claim that an action-desire represents the desired action as good is combined with the Anscombe interpretation, it seems to follow that an action-desire has mind-to-world fit. Recall, according to Anscombean GG theory, an action-desire is correct only if the desired action is good. This suggests that it is the desire, rather than the world, that is incorrect or subject to revision if the desired action is not good. For example, suppose I desire to kick a puppy. Suppose further that it would not be good to kick a puppy. Under such circumstances, we certainly would not wish to say that it is the world—that is, the fact that it would not be good to kick a puppy—that is at fault and should be revised to suit my desire. Rather, it is my desire to kick a puppy that should be revised, given up, or be left unsatisfied. If this is right, then the Anscombe interpretation entails that the desire to kick a puppy has mind-to-world fit.

From the point of view of the Anscombean GG theory, the above conclusion is problematic since it links the formal aim of an action-desire to the fact that it has mind-to-world fit. Recall, according to (II), it is in virtue of the fact that action-desires aim at the good that they may be described as incorrect when the desired action is not good, and it is in virtue of being incorrect when the desired action is not good that an action-desire may be said to have mind-to-world fit. The upshot is that, contra (IV), the aim of an action-desire is linked to the fact that it has mind-to-world fit, rather than to the fact that it has world-to-mind fit. This is a very surprising and unhappy result. One would have expected that the feature in virtue of which an action-desire aims at the good be identified with or closely linked to the feature in virtue of which it has a different direction of fit to belief. This is precisely the view to which (IV) commits the advocate of Anscombean GG theory. But if the present objection is right, the feature in virtue of which an action-desire aims at the good turns out to be exactly the feature in virtue of which an action-desire has the same direction of fit as belief; namely, the fact that an action-desire represents the desired action as good.

3. Background and Definitions

The preceding analysis calls attention to two shortcomings in Anscombean GG theory. First, the Anscombe interpretation of the direction of fit metaphor is unable to accommodate the intuition that attitudes such as supposing and pretending have the same direction of fit as belief. Second, when the Anscombe interpretation is combined with (II), it does not only have the unhappy consequence that an action-desire has mind-to-world fit, but also that it has mind-to-world fit in virtue of the fact that it aims at the good. The upshot is that the claim that action-desires aim at the good cannot satisfy the requirements of (IV), according to which the aim of an action-desire is linked to the fact that it has world-to-mind fit. In light of these difficulties with Anscombean GG theory—and with the Anscombe interpretation of the direction of fit metaphor, in particular—I wish to propose an alternative interpretation of the notion of direction of fit, one that (i) preserves the intuition that the aim of an action-desire is something that explains, is

explained by, or otherwise linked to the fact that it has world-to-mind fit, and (ii) is able to accommodate the intuition that attitudes like supposing and pretending have the same direction of fit as belief. To this end, I claim that attitudes with mind-to-world fit have a different type of content (in a sense soon to be specified) to states with world-to-mind fit. I will refer to this interpretation of the direction-of-fit metaphor as the *two-content interpretation*.

Before we embark on a detailed discussion of the two-content interpretation, it will be helpful to address a few terminological issues. Typically, an assertion is defined as a speech act in which a proposition is represented as being true.[9] It is widely held that an assertion may share the same propositional content as a psychological attitude, such as belief, hope or doubt. For example, the assertion that Goldbach's conjecture is true has the same propositional content as the belief that Goldbach's conjecture is true. Moreover, an agent may use the assertion that Goldbach's conjecture is true to express her belief that Goldbach's conjecture is true. I will refer to an assertion that expresses a belief of the agent making the assertion as a *sincere assertion*. There are cases in which an agent makes a putative assertion, but in which the putative assertion does not express one of the speaker's beliefs. I will refer to such putative assertions as *insincere assertions*. Significantly, to say that a putative assertion is insincere, in the present sense, is not to say that it is somehow infelicitous. Whether an insincere assertion is infelicitous will depend on other factors, such as the context of utterance. For example, if an insincere assertion is made in the context of a theatrical performance, the fact that it does not express a belief of the speaker does not immediately render the assertion infelicitous. However, in a context in which an agent is asked to give her honest opinion on a matter, it would typically be infelicitous for her to make an insincere assertion.

An assertion may be codified in a sentence (written, spoken, or signed) by which the asserted proposition is conveyed. I will refer to particular tokens of such written, spoken, or signed sentences as *utterances*. Two or more different utterances may be used to assert the same proposition. For example, the utterances "It is raining today" and "Es regnet heute" may both be used to assert the proposition *it is raining today*. By necessity, I will be employing written sentences to express the various propositions discussed in this paper. However, it is important to keep in mind that these written sentences will be merely standing proxy for the propositions they are typically used to convey. In order to avoid confusion on this score, I will use quotation marks to indicate when I'm talking about the utterance "It is raining today" and italics to indicate that I am referring to the proposition *it is raining today*.

In addition to assertoric utterances, there are also non-assertoric utterances, such as questions and commands. Assertions, questions, and commands differ in their illocutionary force.[10] However, it is sometimes assumed that utterances with different illocutionary force may have the same propositional content. For example, consider the following three utterances:

[9] If one is an anti-realist about propositions, one may replace talk about propositions with talk about that-clause complements.

[10] See and Cf. Searle and Vanderveken [1985].

(**A**): "The office door is shut."
(**B**): "Is the office door shut?"
(**C**): "Shut the office door!"

It is often held that (A), (B), and (C) share the same propositional content; namely, the proposition: *the office door is shut*.[11] The difference between the three utterances has to do with the illocutionary force with which this single proposition is expressed, with the proposition being asserted in (A), questioned in (B), and commanded in (C). Since, according to this view, it is possible to vary the force of an utterance while keeping the content fixed, it follows that content and force are independent features of an utterance. I will refer to this view as the *independence thesis*.[12]

4. Motivating Illocutionary Content

There are two features of the notion of content presupposed by the independence thesis that makes it less than ideal for our present purposes. First, the independence thesis relies on a notion of content that leaves the logical properties of an utterance underspecified. This is because a complete characterisation of the logical properties of an utterance—including the specification of the various deductive inferences for which a particular utterance may be employed—is at least partly determined by the sentence's illocutionary force. Recall, according to the independence thesis, (A), (B), and (C) all have the same content: the proposition *the office door is shut*. If we assume that this content is sufficient to determine the logical character of all three utterances, then it would follow that (A), (B), and (C) should be logically interchangeable, despite their contrasting illocutionary force. But this does not seem to be the case. For example, the following is a valid deductive inference:

[11] While this is not a view I endorse, it is nevertheless one that I wish to accommodate. However, it is worth noting that, among philosophers who specialise in the semantics of questions, the dominant view is that the content of a question is identical to or to be understood in terms of its answerhood conditions, where "answerhood conditions" is taken to refer to the set of all possible answers, true answers, partial answers, etc. According to this view, which we may refer to as *Hamblin question-semantics*, a question picks out a set of propositions, rather than a single proposition (see: Hamblin [1973] and Karttunen [1977]. For overview of the relevant literature, see: Groenendijk and Stokhof [1997], Higginbotham [1996], and Ginzburg [1996]). Thus, many philosophers would reject the claim that (A) and (B) share the same propositional content. Insofar as such philosophers are committed to saying that (A) and (B) have different types of content, I would consider them allies since, as will soon become clear, I wish to make a similar claim. However, in order to make my account appealing to the widest possible audience, I will attempt to set things up in a way that is neutral between the assumption that (A) and (B) have the same propositional content and Hamblin question-semantics. Similarly, there is a growing number of theorists who reject the claim that (A) and (C) share the same propositional content (e.g., Chellas [1971], Vranas [2008; 2010], Warnock [1976], and Stalley [1972]). I also view such theorists as allies, but will endeavour to remain neutral on this issue here.

[12] For a defense of the independence thesis, see Stenius [1967].

(**A1**): "The office door is shut."
(**A2**): "If the office door is shut, then Professor Smith is away."
(**A3**): "Therefore, Professor Smith is away."

However, neither of the following appears to be a valid deductive inference:

(**B1**): "Is the office door shut?"
(**B2**): "If the office door is shut, then Professor Smith is away."
(**B3**): "Therefore, Professor Smith is away."

and,

(**C1**): "Shut the office door!"
(**C2**): "If the office door is shut, then Professor Smith is away."
(**C3**): "Therefore, Professor Smith is away."

The problem with (B1)-(B3) and (C1)-(C3) is that in both cases, the initial premise fails to satisfy the antecedent of the conditional specified in the second premise ((B2) and (C2), respectively). This leaves us without any basis for inferring the consequent of the conditional, as specified in (B3) and (C3), respectively. Moreover, the failure of (B1) and (C1) to satisfy the antecedent of (B2) and (C2) is directly due to their illocutionary force. (B1) and (C1) both fail to satisfy the antecedent of (B2) and (C2) precisely because they fail to depict the proposition *the office door is shut* as true. In fact, (B1) and (C1) are both perfectly consistent with the falsity of the proposition *the office door is shut*. The same, of course, cannot be said of (A1), which is clearly inconsistent with the falsity of the proposition *the office door is shut*. Thus, although according to the independence thesis, (A1), (B1), and (C1) share the same content, (A1) is logically inconsistent with the negation of the proposition *the office door is shut*, while (B1) and (C1) are both logically consistent with the negation of the proposition *the office door is shut*. Given that the only difference between (A1), (B1), and (C1) has to do with their illocutionary force, it follows that illocutionary force makes a logical difference. The upshot is that if we aspire to have a complete characterisation of the logical properties of an utterance, then we must take into consideration its illocutionary force.

The notion of content which interests me in the present investigation is one that corresponds with a complete characterisation of the logical properties of an utterance—to wit, one that includes illocutionary force. As such, I will not be presupposing the notion of content implicated by the independence thesis. To avoid confusion, and in order to emphasise that the notion of content currently at play is one that includes illocutionary force, I will employ *illocutionary content* as an umbrella term for the notion of content that corresponds with a logically complete characterisation of (A), (B), and (C). There are three distinct types of illocutionary content that are relevant to the present discussion: indicative content (e.g., *the office door is shut*), interrogative content (e.g., *is the office door shut?*), and imperative content (e.g., *shut the office door!*). According to the present view, just as a sincere assertion expresses the attitude of believing, so, too, does a sincere question express the attitude of wondering and a sincere request

expresses an attitude of wanting. Moreover, just as an assertoric utterance may be used to convey indicative content, an interrogative utterance may be used to convey interrogative content and an imperative utterance may be used to convey imperative content. Hence, the illocutionary content of an utterance or attitude corresponds with that which distinguishes a case of believing or asserting from a case of wondering or questioning, and distinguishes both from a case of wanting or requesting.

Having such a notion of content is generally important because we assess an agent's rational standing in light of the illocutionary content of their utterances and attitudes, and not in terms of the propositional content of their utterances and attitudes. For example, if an agent believes or sincerely asserts that the office door is shut, we take them to have or express an attitude that is inconsistent with being agnostic about whether or not the office door is shut. Thus, if an agent were to adopt both an attitude of belief and an attitude of agnosticism towards the office door being shut (assuming that such a combination of doxastic attitudes is possible), we would deem them guilty of irrationality. However, consider the case of an agent who wonders if the office door is shut. Such an agent has an attitude that may be expressed by sincerely asking if the office door is shut. Unlike the agent who believes or sincerely asserts that the office door is shut, the agent who wonders or sincerely asks if the office door is shut has or expresses an attitude that is consistent with being agnostic about the office door being shut. Such an agent would not be deemed irrational for simultaneously having both attitudes. However, according to the independence thesis, wondering if the office door is shut, sincerely asking if the office door is shut, believing that the office door is shut, and sincerely asserting that the office door is shut all have the same propositional content: namely, the proposition *the office door is shut*. Thus, we could not hope to evaluate an agent's rational standing by simply considering the propositional content of their utterances or attitudes. We must look, instead, to the illocutionary content of an agent's utterances and attitudes.

The notion of illocutionary content is also of specific importance to the task of providing an adequate characterisation of the differing directions of fit of beliefs and action-desires. If we conceive of a command as expressing a desire that someone perform a certain action, then we may view an action-desire as akin to a self-issued command that one perform a certain action. On this view, an action-desire has the same illocutionary content as an imperative utterance: namely, imperative content. By contrast, a belief, as we have already observed, has the same illocutionary content as an assertoric utterance: namely, indicative content. In order for an attempt to make sense of the direction of fit metaphor to be adequate, it must not only account for the fact that beliefs and action-desires have different directions of fit, but it must also account for the fact that a belief has the same direction of fit as the attitudes of supposing and pretending.

On this score, I believe that the two-content interpretation is better than the Anscombe interpretation. According to the two-content interpretation, beliefs have the same direction of fit as the attitudes of supposing and pretending because all three have indicative content, while action-desires have the same direction of fit as wishes and

hopes because all three have imperative content. Insofar as indicative content is truth-evaluable, the present view entails that beliefs, supposings, and pretendings all have truth-evaluable illocutionary content. However, what sets the attitude of belief apart from the attitudes of supposing and pretending is that a belief is correct only if its content is true. This is, in part, what it means to say that belief aims at truth. On the present account, since it is possible for a supposing or pretending to represent matters in a way that they are not, it is also possible for their illocutionary content to be false. However, it does not follow from the fact that the illocutionary content of a supposing or pretending is false that either attitude is incorrect or subject to revision. This is at least part of what it means to say that supposing and pretending do not aim at truth. Hence, the present framework is able to preserve the distinction between saying that belief has a certain direction of fit and saying that belief has a certain aim.

Just as there is widespread agreement that attitudes such as hypothesising, supposing, and pretending share the same direction of fit as belief, it is also relatively uncontroversial to hold that other attitudes, such as wishing and hoping, share the same direction of fit as desire. However, it is not obvious that there is a sharp distinction between our quotidian conception of a desire and our quotidian conception of a wish. In fact, there are contexts in which the terms are treated as synonymous. For example, saying that one wishes to visit Spain and that one desires to visit Spain may both be taken to express the same attitude. Nevertheless, there are also subtle differences in how we typically employ the two terms. For example, consider the following pair of self-descriptions:

(**D1**): "I wish I were never born."
(**D2**): "I desire I were never born."

While (D1) seems perfectly natural, there is something awkward about (D2). It is plausible that the awkwardness of (D2) is in part due to the fact that it is typical to employ the word "wish" when describing one's attitude to something one recognises to be unattainable.

While perhaps suggestive, the preceding observation falls short of establishing a non-overlapping distinction between our quotidian conception of a wish and a desire. However, we may arrive at a non-overlapping distinction between a wish and an action-desire by imposing the following theoretically-motivated restriction on both terms. I will use the expressions "desire" and "action-desire" to refer to conative attitudes that are directed at something an agent believes to be a possible future outcome.[13] On this

[13] Velleman seems to commit himself to a similar conception of desire when he notes that "one cannot desire something if it seems impossible or if it seems already to have come about; one can desire that *p* only if *p* seems attainable, in the sense of being a possible future outcome"[1992: 17]. However, Velleman describes his picture of desires as a "hypothesis". I find this odd, for it gives the impression that his account is supposed to be tested against some observable fact. That would, of course, be an appropriate characterisation if one were making an empirical claim to the effect that there is some neurological state to which the term "desire" is applicable, and an agent can only enter into this neurological state with respect to something she believes to be a possible future outcome. However, this strikes me as highly implausible.

view, one may desire to have a glass of wine (assuming that one believes that having a glass of wine is a possible future outcome), but one may not desire that one were never born (assuming that one believes that undoing one's birth is not a possible future outcome). By contrast, I will be using the term "wish" to refer to a conative attitude that is directed at something an agent believes not to be a possible future outcome.

With the preceding terminological distinction in place, the two-content interpretation allows us to consistently hold that action-desires and wishes have the same direction of fit, and yet deny that they share the same aim or correctness conditions. To this end, I hold that action-desires, wishes, and hopes all have imperative content. Moreover, I maintain that imperative content has a two-valued logical structure analogous to that of indicative content. However, while in the case of indicative content the relevant values are "true" and "false", in the case of imperative content the relevant values are something like "attainable" and "unattainable". Following Velleman, I define the attainable as a "possible future outcome".[14] Thus, I hold that the illocutionary content of action-desires, wishes and hopes all have the same two-valued logical structure such that the thing desired, wished, or hoped for is either a possible future outcome or not a possible future outcome. However, one of the things that set an action-desire apart from a wish is that an action-desire is correct only if its object is attainable. Thus, contra Anscombean GG theory, I hold that the aim of an action-desire—alluded to in (IV)—is the attainable.

5. Formal Presuppositions and Illocutionary Implications

At the end of the previous section, I suggest that what distinguishes an action-desire from a wish is that the former is correct only if its object is attainable. However, this introduces a puzzle. How can we consistently affirm that an action-desire is correct only if the desired action is attainable and yet deny that an action-desire represents its object as attainable? After all, if one were to hold that an action-desire represents its object as attainable, then there is an obvious and intuitive sense in which an action-desire gets things wrong if the desired action is not attainable. Simply put, the desire represents matters incorrectly. However, if one holds that an action-desire does not represent its object as attainable, then it seems inappropriate to say that the desire gets things wrong if the desired action is not attainable. That would be analogous to criticising someone for making a claim that they never made, or at least so the present worry goes.

My strategy for solving this puzzle is to argue that an action-desire presupposes (in a sense to be specified momentarily) that its object is attainable, and may therefore be

---

Alternatively, the claim that desires aim at the attainable may be seen as a hypothesis about certain socio-linguistic facts regarding how we use the term "desire". But it is certainly false, as a socio-linguistic claim, that we only use the word desire to refer to conative attitudes an agent takes towards things she believes to be attainable. Moreover, it is unclear what else Velleman could mean when he describes his proposed conception of a desire as a hypothesis. By contrast, I offer the present account of action-desires, not as a hypothesis, but as a stipulation; that is to say, I will be applying the label "action-desires" to the set consisting of conative attitudes an agent takes towards actions she believes she can perform.

[14] Velleman [1992: 17].

described as correct only if its object is attainable. In non-technical terms, a presupposition is defined as information that is assumed or taken for granted by an individual or utterance.[15] For example, suppose I were to make the following sincere assertion:

(**E1**): "Usain Bolt has won the 100 meters once again."

In so doing, I may plausibly be said to presuppose that the following propositions are true:

(**E2**) *There is a (salient and identifiable) Usain Bolt.*
(**E3**) *There is a (salient and identifiable) 100 meters.*
(**E4**) *Usain Bolt has won the 100 meters at least once before.*

The technical notion of a presupposition that I wish to introduce here aims to mirror the results of the non-technical notion. Thus, I wish to preserve the intuition that (E1) presupposes (E2), (E3), and (E4). However, the technical notion derives much of its theoretical motivation from Frege's observation "[t]hat the name 'Kepler' denotes something is just as much a presupposition for the assertion 'Kepler died in misery' as for the contrary assertion."[16] Within Frege's framework, a presupposition is conceived of as a definedness condition, a requirement for an expression to have a truth-value. Strawson takes up this idea by defining a presupposition along the following lines:

Definition 1 (Strawsonian Presupposition):
ϕ presupposes ψ *IFF* ψ is true whenever ϕ is true or false.[17]

I do not wish to follow Frege-Strawson by describing a presupposition as a prerequisite for an utterance to have a truth-value. However, what I wish to take from Frege-Strawson is the idea that if the truth of a certain proposition is presupposed (rather than asserted) by an utterance, then it should also be implied (in some sense) by the negation of that assertion.

[15] Describing a presupposition as an "assumption" is potentially controversial, if by "assumption" we mean the assumption of an agent. Theories of presupposition may be divided into *pragmatic* and *semantic*. According to the standard pragmatic theory, presuppositions are tied to the attitudes of a particular agent. For example, according to Stalnaker [1974], an utterance presupposes a proposition because the speaker took the proposition for granted when she made the utterance. By contrast, according to the standard semantic theory, presuppositions are tied to sentences, rather than to the attitudes of an agent. For example, according to one interpretation of Strawson [1950], an utterance presupposes a proposition because the truth of the proposition is a necessary condition for the utterance to denote, be meaningful, or have a truth-value. In the technical account of a presupposition that follows, I wish to remain non-committal on the pragmatic/semantic debate. As such, I will not appeal to the notion of an "assumption" in my formulation of the technical notion of a presupposition.
[16] Frege [1892: 168].
[17] Strawson [1952]. For a discussion of Strawson's proposals, see: Sellers [1954]; Strawson [1954]; and Nerlich [1967].

The motivating intuition behind my borrowing from Frege-Strawson may be summarised as follows. On the one hand, saying that an utterance, ϕ, presupposes some proposition, ψ, means that an agent who utters ϕ can be expected to believe ψ. This suggests that if ψ is presupposed by ϕ, then ψ should be implied (in some sense) by ϕ. On the other hand, insofar as a presupposed proposition is not asserted by an utterance, then accepting the proposition(s) asserted by an utterance should not be a prerequisite for accepting a presupposed proposition. This means that whatever property of an utterance in virtue of which it implies a presupposed proposition should be equally effective in giving rise to the relevant implication even if one were to reject all the propositions that the utterance asserts. Thus, the implication should also hold if one were to make an utterance that was like the original utterance in every respect except in that all of the propositions asserted in the original utterance were now denied, otherwise one would be justified in concluding that the proposition was a member of the set of propositions asserted by the original utterance, in which case it would fail to qualify as a presupposed proposition. The upshot is that ψ may be said to be presupposed by ϕ if and only if ϕ implies ψ and $\neg\phi$ implies ψ. I will refer to this technical conception of a presupposition as a *formal presupposition*.[18]

However, it is important to get clear on the notion of implication currently at play. If we model the present notion of an "implication" on classic semantic entailment, then ϕ implies ψ just in case it is impossible for ϕ to be true and ψ to be false. However, since I will be considering utterances and attitudes that are not truth-evaluable (or whose illocutionary content is not truth-evaluable), the classic notion of semantic entailment will not do for present purposes. There is also a much more general problem with modelling the relevant notion of an implication on classic semantic entailment. According to classic semantic entailment, if ϕ implies ψ and $\neg\phi$ implies ψ, then it follows that ψ is a tautology. This is a very unhappy result since it would follow from the definition of a formal presupposition that all formal presuppositions are tautologies.

In order to avoid the above difficulties, I wish to appeal to a notion of implication that is both broader and weaker than classic semantic entailment. I will refer to this notion of implication as an illocutionary implication (***i***-implication), which I define as follows:

[18] Something along the lines of this definition of a presupposition is advanced by Bas van Fraassen [1968]. Survival under negation is widely regarded as the most important diagnostic test of a presupposition. A presupposition that passes the negation-embedding test is said to be "projected" under negation-embedding. However, projection under negation-embedding is typically seen as falling short of a necessary condition for something to count as a presupposition because there are circumstances in which presuppositions fail to project. Indeed, attempting to explain why presuppositions sometimes fail to project is one of the central problems in presupposition theory. My account of presupposition departs from the standard picture since it takes passing the negation-embedding test to be both necessary and sufficient for a proposition to be presupposed by an utterance or attitude. This is largely due to the fact that the current technical account of a presupposition does not exploit the notion of "projection" that exercises so much of presupposition theory. For an excellent overview of presupposition literature, see Soames [1989].

Definition 2 (Illocutionary Implication):
ϕ ***i***-implies ψ *IFF* an utterance of ϕ indicates that the speaker believes ψ.[19]

Illocutionary implication is broader than classic semantic entailment since it has application not only to utterances with truth-conditions (i.e., ones with indicative content), but also to utterances with attainability-conditions (i.e., ones with imperative content). It is also weaker than the classic semantic entailment since it attempts to track, not when a presupposition is entailed by a particular utterance, but rather what an utterance indicates that the speaker believes.[20] With the broader notion of an ***i***-implication in hand, we may now define a formal presupposition as follows:

Definition 3 (Formal Presupposition):
ϕ formally presupposes ψ *IFF* ϕ ***i***-implies ψ and $\neg\phi$ ***i***-implies ψ.

Definition 3 provides us with a diagnostic test for determining if a particular proposition is formally presupposed by an utterance. On the current view, ϕ formally presupposes ψ just in case ϕ ***i***-implies ψ and there is some μ such that μ is formed by embedding ϕ under a negation operator, and μ ***i***-implies ψ. Hence, we may determine if a proposition that is ***i***-implied by an utterance is formally presupposed by that utterance by first embedding the illocutionary content of the utterance under a negation operator, and then checking to see if the ***i***-implication still holds. Call this the *negation-embedding test*. For example, if we embed (E1) under the negation operator we arrive at (E1*):

(**E1***) "Usain Bolt has not won the 100 meters once again."

According to the negation-embedding test, if (E2)-(E4) are ***i***-implied by both (E1) and (E1*), then (E2)-(E4) are formally presupposed by (E1). An examination of (E1) and (E1*) reveals that both ***i***-imply (E2)-(E4). That is to say, an utterance of (E1), as well as an utterance of (E1*), indicate that the speaker believes (E2), (E3), and (E4). Consequently, (E2), (E3), and (E4) all pass the negation-embedding test. It follows that (E2)-(E4) are formal presuppositions of (E1). Contrast this with (E5):

(**E5**) *Usain Bolt did not lose the 100 meters.*

It is clear that (E1*) does not ***i***-imply (E5), since an utterance of (E1*) would not indicate that the speaker believes (E5). Consequently, (E5) fails the negation-embedding test; it is not a formal presupposition of (E1).

[19] See and Cf. Beaver [2001: 15].
[20] The notion of an utterance indicating what the speaker means is here deliberately left vague in order to remain neutral between semantic and pragmatic conceptions of a presupposition. The reader is free to unpack the notion in terms of her preferred theoretical account.

6. Action-Desires, the Attainable, and the Good

The first step towards solving the puzzle of how an action-desire may be correct only if the desired action is attainable, even though an action-desire does not represent its object as attainable, is to register that imperative utterances also seem to have presuppositions. For example, consider the following imperative utterance:

(**F1**): "Shut the office door!"

(F1) has the following putative presuppositions:

(**F2**): *There is an (salient and identifiable) office door*.
(**F3**): *The office door is open*.

It should be obvious that an utterance of (F1) indicates that the speaker believes (F2) and (F3). We may therefore conclude that (F1) ***i***-implies (F2) and (F3). However, it is not immediately clear that (F1) is a candidate for the negation-embedding test. After all, I have been insisting that imperative content is not truth-evaluable. How then are we to characterise the negation of a non-truth-evaluable item? In response to this problem, I propose that we conceive of the negation of, for instance, a command to φ in terms of a command to refrain from φ-ing. Hence, we may form the negation of (F1) by embedding it under a "do not" operator, in order to form (F1*):

(**F1***): "Do not shut the office door!"

We may now ask if (F1*) ***i***-implies (F2) and (F3). That is to say, would an utterance of (F1*) indicate that the speaker believes (F2) and (F3)? Since it would be pointless to command someone not to shut the office door if there were no (salient and identifiable) office door and if the office door were not open, an utterance of (F1*) does indicate that the speaker believes (F2) and (F3). Consequently, (F2) and (F3) both pass the negation-embedding test for imperative utterances, and are therefore formal presuppositions of (F1).

With the theoretical apparatus for determining the formal presuppositions of an imperative utterance now in hand, we may ask if the command to shut the office door formally presupposes that shutting the office door is attainable. In response to this question, I submit that the set of propositions that must be true in order for shutting the office door to be attainable is always a subset of the set of propositions formally presupposed by the command to shut the office door. Or, to put the same point more generally and in terms of a conditional, if a certain proposition must be true in order for the object of an imperative utterance to be attainable, then that imperative utterance formally presupposes that proposition. Let ψ stand for a proposition, ϕ stand for an imperative utterance, and ¬ϕ stand for the imperative utterance formed by embedding ϕ under the "do not" operator. The claim that an imperative utterance always formally

presupposes a proposition that must be true in order for the imperative to be attainable may be demonstrated by the following argument:

(1) If ψ must be true for ϕ to be attainable, then ϕ ***i***-implies ψ. (Premise defended below)

(2) If ψ must be true for ϕ to be attainable, then ¬ϕ ***i***-implies ψ. (Premise defended below)

(3) If ψ must be true for ϕ to be attainable, then ϕ ***i***-implies ψ and ¬ϕ ***i***-implies ψ. (From (1) and (2))

(4) If ϕ ***i***-implies ψ and ¬ϕ ***i***-implies ψ, then ϕ formally presupposes ψ. (Definition of a Formal Presupposition)

(5) If ψ must be true for ϕ to be attainable, then ϕ formally presupposes ψ. (From (3) and (4))

Premise (1) rests on the intuition that it would be pointless to command someone to perform an action they could not perform. As such, a sincere imperative utterance would indicate that the speaker believes that the imperative is attainable. Since A ***i***-implies B just in case an utterance of A would indicate that the speaker believed B, a sincere imperative utterance always ***i***-implies that the imperative is attainable. Premise (2) rests on the intuition that it would be pointless to command someone to refrain from performing an action that they could not perform. For example, it would be pointless to issue the command, "Do not shut the office door!" if it was not possible to shut the office door. If this intuition is right, then ¬ϕ always ***i***-implies those propositions that must be true in order for ϕ to be attainable. The other premises of the argument follow from (1), (2), and the definition of a formal presupposition, as specified in (4). We may therefore conclude that the set of propositions that must be true for an imperative utterance to be attainable is always as subset of the set of propositions that are formally presupposed by that imperative utterance.

Let us conclude the present section by considering the question of whether an action-desire formally presupposes that the desired action is good. If the GG theorist is going to claim that, for example, the command to shut the office door formally presupposes that shutting the office door is good, then she will need to show that (F4), or some comparable proposition, passes the negation-embedding test for (F1):

(**F4**): It is good to shut the office door.

However, in order to pass the negation-embedding test for (F1), (F4) must be ***i***-implied by both (F1) and (F1*).

(**F1**): "Shut the office door!"
(**F1***): "Do not shut the office door!"

But this is clearly not the case. Even if a sincere utterance of (F1) indicates that the speaker believes (F4)—a claim that strikes me as dubious—a sincere utterance of (F1*) certainly does not indicate that the speaker believes (F4). Therefore, it follows from the definition of an ***i***-implication that (F1*) does not ***i***-imply (F4). (F4) therefore fails the negation-embedding test for (F1).

7. Conclusion

In this paper, I have accomplished two things. First, I have argued that the two-content interpretation is superior to the Anscombe interpretation because it allows us to preserve the distinction between saying that a psychological attitude has a certain direction of fit and saying that it has certain correctness conditions. According to the two-content interpretation, desires and beliefs have different directions of fit because they have different kinds of illocutionary content: indicative content in the case of the former, and imperative content in the case of the latter. However, a belief is correct only if it is true, not because it has indicative content, but because it aims at the truth.

Second, I have argued that if we assume that the aim of an action-desire is linked to the fact that it has world-to-mind fit, then it is more plausible that desires aim at the attainable than at the good. Since, according to Anscombean GG theory, an action-desire aims at the good in virtue of representing the desired action as good, and since a desire represents the desired action as good in virtue of having mind-to-world fit, it follows that an action-desire aims at the good in virtue of having mind-to-world fit. However, according to the view defended in this paper, an action-desire aims at the good because its imperative content formally presupposes that the desired action is attainable. Since an action-desire has world-to-mind fit in virtue of having imperative content, the fact that an action-desire aims at the attainable is linked to the fact that it has world-to-mind fit.

References:

Anscombe, G. E. [2000], *Intention*. Cambridge, MA: Harvard University Press.

Beaver, D. [2001], *Presupposition and Assertion in Dynamic Semantics*. (Stanford, CA: CSLI Publications).

Byrne, A. and Hájek, A. [1997], "David Hume, David Lewis, and Decision Theory", *Mind*, **106**: 411-28.

Chellas, B. F. [1971], "Imperatives", *Theoria*, **37**: 114-129.

Davidson, D. [1984a], "Moods and Performances", *Inquiries into Truth and Interpretation*. (Oxford: Clarendon Press): 109-122.

Davidson, D. [1984b], "Communication and Convention", *Inquiries into Truth and Interpretation*. (Oxford: Clarendon Press): 265-280.

Engel, P., [2004], 'Truth and the Aim of Belief', in D. Gillies (ed.), *Laws and Models in Science*. (London: King's College Publications).

Fraassen, B. C. [1968], "Presupposition, Implication, and Self-Reference", The Journal of Philosophy, **65** (5): 136-152.

Grice, H. P. [1957], "Meaning", *The Philosophical Review*, **66** (3): 337-388.

Nerlich, G. [1967], "Presupposition and Classical Logical Relations", *Analysis*, **27** (3): 104-106.

Price, H. [1989], "Defending Desire-as-Belief", Mind, **98**: 119-127.

Stalley, R. F. [1972], "Intentions, Beliefs, and Imperative Logic", *Mind*, **81**: 18-28.

Searle, J. R. [1985], *Expression and Meaning: Studies in the Theory of Speech Acts*, New York: Cambridge University Press.

Searle, J. R. and Vanderveken, D. [1985], *Foundations of Illocutionary Logic*. (Cambridge: Cambridge University Press).

Sellars, W. [1954], "Presupposing", *Philosophical Review*, **63**: 197-215.

Smith, M. [1987], "The Humean Theory of Motivation", *Mind*, **381**: 36-61.

Soames, S. [1989], "Presupposition", in Gabbay, D. and Guenther, F. (eds.), *Handbook of Philosophical Logic*. (Dordrecht: Reidel), **4**: 553-616.

Stalnaker, R. [1974], "Pragmatic Presuppositions", in Munitz, M. and Unger, P. (eds.), *Semantics and Philosophy*. (New York: New York University Press): 197-214.

Stampe, D. [1987], "The Authority of Desire", *Philosophical Review*, **96**: 335-81.

Stenius, E. [1967], "Mood and Language-Game", *Synthese*, **17**: 254-274.

Strawson, P. F. [1950], "On Referring", *Mind*, **59**: 320-344.

Strawson, P. F. [1952], *Introduction to Logical Theory*. (London: Methuen).

Strawson, P. F. [1954], "A Reply to Mr. Sellars", *Philosophical Review*, **63**: 216-231.

Tenenbaum, S. [2007], *Appearances of the Good: An Essay on the Nature of Practical Reason*. (Cambridge, UK: Cambridge University Press).

Tenenbaum, S. [2010], *Desire, Practical Reason, and the Good*. (Oxford: Oxford University Press).

Velleman, D. [1992], "The Guise of the Good", *Nous*, **26** (1): 3-26.

Vranas, P. B. M. [2008], "New Foundations for imperative Logic I: Logical Connectives, Consistency, and Quantifiers", *Nous*, **42**: 529-572.

Vranas, P. B. M. [2010], "In Defense of Imperative Inference", *Journal of Philosophical Logic*, **39**: 59-71.

Warnock, G. [1976], "Imperatives and Meaning," in Lewis, H. D. (ed.), *Contemporary British Philosophy: Personal Statements*. (London: Allen and Unwin), **4**: 294-304.