Reconceiving Direction of Fit

Avery Archer

ABSTRACT: I argue that the concept of direction of fit is best seen as picking out a certain inferential property of a psychological attitude. The property in question is one that believing shares with assuming and fantasizing and fails to share with desire. Unfortunately, the standard analysis of direction of fit obscures this fact because it conflates two very different properties of an attitude: that in virtue of which it displays a certain direction of fit, and that in virtue of which it displays certain revision-conditions. I claim that the latter corresponds with the aim of an attitude, not its direction of fit. In order to remedy this failure of the standard analysis, I offer an alternative account of direction of fit, which I refer to as the two-content analysis.

KEY WORDS: Direction of Fit, Theoretical Attitudes, Practical Attitudes, Belief, Desire

1. Introduction

What property of an attitude is the notion of direction of fit (henceforth, DOF) meant to capture? This is the question I take up in this paper. The answer I defend is that DOF picks out a certain inferential property of an attitude. I argue that the property in question is one that believing has in common with assuming and fantasizing, and which it fails to share with desire. Unfortunately, the standard analysis of DOF obscures these observations because it conflates two very different properties of an attitude: that in virtue of which it displays a certain DOF, and that in virtue of which it displays certain revision-conditions. In order to remedy this failure of the standard analysis, I offer an alternative account of DOF, which I call the *two-content analysis*.

2. The Standard Analysis

Mark Platts summarises the central intuition behind the direction of fit analysis of psychological attitudes as follows:

> The distinction is in terms of direction of fit of mental states to the world. Beliefs aim at being true, and their being true is their fitting the world; falsity is a decisive failing in a belief, and false beliefs should be discarded; beliefs should be changed to fit the world, not vice versa. Desires aim at realization, and their realization is the world fitting with them; the fact that the indicative content of a desire is not realised in the world is not yet a failing in the desire, and not yet any reason to discard the desire; the world, crudely, should be changed to fit with our desires, not vice versa.[1]

---

[1] Platts (1979, p. 257). See: Price (1989) and Humberstone (1992) for discussion.

Platts unpacks DOF in terms of whether the mind or the world is subject to revision when there is a lack of correspondence between the two.[2] According to this approach, saying that a belief has mind-to-world DOF entails that when the world and the belief fail to line up, the burden is on the believing agent to alter her mental representations so as to bring them into line with the world. By contrast, to say that a desire has world-to-mind DOF entails that when the world and the desire fail to line up, the burden on the desiring agent is to alter the world so as to bring it into line with her mental representations. I will refer to the preceding analysis of the DOF metaphor as the *standard analysis*.[3]

The contrast between mind-to-world and world-to-mind DOF is often taken to correspond with the contrast between *cognitive* and *conative* attitudes, respectively (to use David Velleman's preferred terminology)[4], or the contrast between *thetic* and *telic* attitudes, respectively (to use Lloyd Humberstone's preferred terminology)[5], or the contrast between *theoretical* and *practical* attitudes, respectively (to use my preferred terminology). In the discussion that follows I will be presupposing this correspondence. Henceforth, I will use the label "theoretical" to refer to belief-like attitudes, and the label "practical" to refer to desire-like attitudes. This is purely for the sake of convenience. Most people find the labels "theoretical" and "practical" easier to keep track of than the traditional "mind-to-world" and "world-to-mind" locutions. However, all of my arguments may be restated, mutatis mutandis, in terms of the cognitive/conative and thetic/telic dichotomies.

I believe that the standard analysis of DOF is inadequate because, inter alia, it is unable to preserve the intuition that certain attitudes—such as assuming and fantasizing—share the same DOF as belief. The notion of DOF is most commonly discussed in the context of a contrast between belief and desire. However, it is widely held that the distinction is not restricted to belief and desire alone.[6] To say that an attitude has a particular DOF is to conceive of it as a member of a broad class of attitudes with which it shares some essential feature. David Velleman puts the point as follows:

> There are many cognitive attitudes other than belief, attitudes that have the same direction of fit and consequently take the same constitutive predicate. Hypothesizing that **P**, assuming that **P**, fantasizing that **P**, and the like are all attitudes in which **P** is regarded, not as a representation of what is to be brought about, but rather as a representation of what is.[7]

In the cited passage, Velleman observes that several attitudes, apart from belief, display mind-to-world DOF, attitudes that include hypothesizing, assuming, and fantasizing. However, while believing that **P** and hypothesising that **P** are plausibly thought of as subject to revision if

---

[2] See and Cf. Anscombe (2000, p. 56).

[3] Something along the lines of what I am calling the standard analysis appears in Platts (1979), Searle (1983), Price (1989) and Moran (2004). The chief inspiration for the standard analysis comes from Anscombe (2000). However, it remains a matter of some debate whether Anscombe herself should be considered a DOF theorist (See Frost [2014]). For an alternative to the standard analysis, see Smith (1987) and Humberstone (1992).

[4] Velleman (1992).

[5] Humberstone (1992).

[6] For examples of direction of fit being applied to attitudes other than belief and desire, see: Searle (1983, p. 88); Velleman (1992, p. 12); Moran (2004, p. 46-47); Tuomela (2007, p. 277), and Camp (2009).

[7] Velleman (1992, p. 12).

it is false that **P**, the same cannot be said of the attitudes of assuming that **P** and fantasizing that **P**. Although assuming that **P** and fantasizing that **P** represent things as being a certain way, neither is necessarily subject to revision if **P** is false. Consequently, if we assume, in keeping with the standard analysis, that an attitude is theoretical only if it is subject to revision when the attitude and the world fail to line up, then we seem forced to deny that assuming and fantasizing are theoretical attitudes.[8]

The preceding observation suggests a possible diagnosis of where the standard analysis goes wrong. To say that a particular attitude is theoretical is not yet to say anything about when that attitude is subject to revision. Whether or not an attitude is subject to revision seems tied to its aim, and not to its DOF. Hence, a belief is subject to revision if the proposition believed is false because belief aims at truth.[9] Since the attitudes of assuming and fantasizing, by contrast, do not aim at the truth, they are not subject to revision if the proposition assumed or fantasized is false. In light of this, we need to make sense of the DOF metaphor in a way that preserves the distinction between saying that an attitude has a certain aim (and by extension, certain revision-conditions) and the claim that an attitude has a particular DOF. Unfortunately, the standard analysis falls short in precisely this respect.

3. Illocutionary Content

In light of the aforementioned shortcoming in the standard analysis, I wish to propose an alternative analysis of DOF. Here, in brief, is my proposal: an attitude is theoretical just in case it displays the same inferential properties as an indicative speech-act and practical just in case it displays the same inferential properties as an imperative speech-act. I will unpack this proposal at present. I begin with the contrast between indicative and imperative speech-acts. Consider the following pair of utterances:

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

Typically, someone who utters (A) is performing an indicative (or declarative) speech-act, and someone who utters (B) is performing an imperative (or directive) speech-act. Intuitively, (A) and (B) share something in common. They are both, in some sense, *about* the office door and its being shut. Some philosophers have attempted to capture this intuition by suggesting that what (A) and (B) have in common is their propositional content; namely, the proposition: [the office door is shut].[10] The difference between the two utterances has to do with the illocutionary

---

[8] One response to the present objection would be to switch from an analysis of DOF that focuses on revision-conditions to one that focuses on correctness-conditions. This has its advantages. While a premise assumed for reductio may not be subject to revision, it may still be deemed incorrect. However, this strategy does not work in the case of fantasizing that **P** since fantasizing that **P** is not plausibly thought of as incorrect if **P** is false.

[9] While there have been challenges to the claim that belief aims at truth (see Owens [2003]), it remains the orthodox view among philosophers. For a helpful overview of the topic, see Engel (2004) and Chan (2013).

[10] Here and henceforth, I use double quotation marks to indicate when I'm referring to a particular token utterance, and square brackets to indicate when I'm referring to the abstract proposition that constitutes the content of an utterance.

force with which this single proposition is expressed, with the proposition being asserted in (A) and commanded in (B). 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*.[11]

It is worth noting that the independence thesis is not free from controversy. Indeed, while I do believe that (A) and (B) hold something in common, I am sceptical about the thesis that what they have in common is a proposition—i.e., something truth-evaluable. Hence, I am disinclined to accept the independence thesis. Even so, the discussion below will assume (if only for the sake of argument) that the independence thesis is true. My motivation for this is twofold. First, something like the independence thesis continues to enjoy widespread acceptance among philosophers and may even be considered the orthodox view.[12] Second, presupposing the truth of the independence thesis will make defending the two-content analysis more difficult. Hence, if I am able to mount a successful defence of the two-content analysis given the assumption that the independence thesis is true, then (a fortiori) I should be able to offer a successful defence of it when the truth of the independence thesis is not assumed.

I take an inference to be a type of psychological transition (e.g., the transition from one belief or set of beliefs to another), which we may encode in a speech-act that allows us to represent the psychological transition to others and ourselves. I will refer to whatever feature of a speech-act determines the kinds of inferences it may be used to encode its inferential character. Moreover, we may encode an inference in a speech-act by which Significantly, the independence thesis relies on a notion of content that leaves the inferential properties of a speech-act—i.e., the kinds of speech-acdeductive inferences a particular speech-act may be used to perform—unspecified. This is because the inferential properties of a speech-act is at least partly determined by its illocutionary force. Recall, according to the independence thesis, (A) and (B) both have the same propositional content: the proposition [the office door is shut]. If we assume that this content is sufficient to determine their inferential properties, then it would follow that (A) and (B) should be logically interchangeable, despite their contrasting illocutionary force. But this is clearly not the case. For example, an agent who made the following sincere assertions is performing a valid inference:

> (**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, the following putative inference is not valid:

> (**B1**): "Shut the office door!"
> (**B2**): "If the office door is shut, then professor Smith is away."
> (**B3**): "Therefore, professor Smith is away."

---

[11] See and cf. Frege (1879; 1979); Searle & Vanderveken (1985, p. 1).
[12] For a defence of the independence thesis, see: Stenius (1967). For criticisms of the independence thesis, see: Mastop (2011).

The problem with (B1)-(B3) is that the initial premise fails to satisfy the antecedent of the conditional specified in (B2). This leaves us without any basis for inferring the consequent of the conditional, as specified in (B3). Moreover, the failure of (B1) to satisfy the antecedent of (B2) is directly due to its illocutionary force. (B1) fails to satisfy the antecedent of (B2) precisely because it fails to depict the proposition [the office door is shut] as true. In fact, (B1) is 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]. Although, according to the independence thesis, (A1) and (B1) share the same content, (A1) is logically inconsistent with the negation of the proposition [the office door is shut], while (B1) is logically consistent with the negation of the proposition [the office door is shut]. Given that the only difference between (A1) and (B1) has to do with their illocutionary force, it follows that illocutionary force makes a difference with regards to the kinds of valid inferences a speech-act may be used to perform.

The takeaway of the preceding discussion is that if we are interested in the inferential properties of a speech-act, we cannot simply consider its propositional content. We must also consider its illocutionary force. Given that the inferential properties of a speech-act is determined by its propositional content and illocutionary force, it will be helpful to have a label for the combination of both features of a speech-act. I will use the label "illocutionary content" for this purpose. In short, the *illocutionary content* of an utterance is the combination of its illocutionary force and propositional content. Hence, we may now say that the types of deductive inferences a speech-act may be used to perform is determined by its illocutionary content.

4. Psychological Content

While a speech-act is typically conceived of as involving a propositional content and illocutionary force, an attitude is typically conceived of as involving a propositional content and *psychological mode* of presentation.[13] I will use the label "psychological content" to refer to the combination of the propositional content and psychological mode of an attitude. In short, psychological content is to an attitude what illocutionary content is to a speech-act. Moreover, just as we may distinguish between speech-acts with indicative and imperative illocutionary contents, I claim that we may distinguish between attitudes with indicative and imperative psychological contents. According to the present suggestion, if I believe that Hermione shut the door, then my belief displays indicative psychological content, and if I desire that Hermione shut the door, then my desire displays imperative psychological content. Hence, I hold that there is a type of parity between indicative speech-acts (e.g., assertions) and attitudes with indicative psychological content (e.g., beliefs), and a type of parity between imperative speech-acts (e.g., commands) and attitudes with imperative psychological content (e.g., desires). Indeed, we often express beliefs (such as the belief that Hermione shut the door) via an assertion (such as the assertion: "Hermione shut the door.") and we often express desires (such

---

[13] See and Cf: Searle (1983:12ff.) and Kissine (2009: 126).

as the desire that Hermione shut the door) via a command (such as the command, "Hermione, shut the door!").

Assuming that my analogy between psychological and illocutionary content is apt, it follows that the logical properties of an attitude—i.e., that which determines the kinds of deductive inferences in which an attitude may feature—is determined not by the attitude's propositional content, but by its psychological content. For example, one can validly infer from the belief that the door is closed and the belief that if the door is closed, then professor Smith is away that professor Smith is away. However, the same is not true of the desire to close the door, even though (according to the independence thesis) the belief and desire share the same propositional content. Hence, if we are interested in the types of deductive inferences in which an attitude may validly feature, then we must pay attention to the attitude's psychological content, rather than its propositional content.

5. The Two Content Analysis

I believe the claim that an attitude displays a certain DOF is a claim about the psychological content of that attitude. For example, to describe believing as a theoretical attitude is to say that it displays indicative psychological content, and to describe desiring as a practical attitude is to say that it displays imperative psychological content. According to the present view, what all theoretical attitudes—such as believing, assuming, and fantasizing—have in common is that they all display indicative psychological content. Moreover, I hold that an attitude displays indicative psychological content just in case its psychological content is truth-evaluable. In other words, an attitude displays indicative psychological content just in case the combination of its propositional content and psychological mode yields something truth-evaluable.

I now come to a question that may have been troubling some readers. What do we gain from conceiving of assuming and fantasizing as theoretical attitudes? Given that neither is subject to revision when their content (propositional or otherwise) is false, why should we hold that they have truth-evaluable psychological content? Why not think that assumings and fantasizings, like desires, are practical attitudes? The answer is that, like belief, both assumings and fantasizings can feature in *classically valid* inferences. This is obvious in the case of assumptions, since we often use assumptions in the premises of arguments. However, it is true of fantasizing as well. For example, suppose that I am pretending to be Harry Potter (say, while attending a science fiction and fantasy convention), and that my fantasy prompts me to make the following set of utterances:

(**C1**): "I am Harry Potter."
(**C2**): "If I am Harry Potter, then Voldemort killed my parents."
(**C3**): "Therefore, Voldemort killed my parents."

By stipulation, (C1) and (C3) are not things I believe. I do not believe I am Harry Potter; nor do I believe that Voldemort killed my parents. However, (C1)-(C3) remains a perfectly respectable deductive inference for all that. This is because (C1)-(C3) conforms to the classical conception of validity; the truth of its premises guarantees the truth of its conclusion. In sum, believing,

assuming, and fantasizing are indicative attitudes because they all share the same psychological content, and they all share the same psychological content because they may all feature in classically valid inferences: inferences in which the truth of the premises guarantees the truth of the conclusion.

Practical attitudes, by contrast, cannot feature in classically valid inferences. This follows from the conjunction of the following two claims: (i) the psychological content of practical attitudes is not truth-evaluable and (ii) the classical conception of validity only applies to truth-evaluable items. As such, practical attitudes cannot feature in deductive inferences that involve truth-evaluable premises and conclusions (i.e., classically valid inferences).

It is worth emphasising that the present line of argument does not assume that practical attitudes cannot feature in inferences of their own. For example, the following also appears to be an acceptable inference:

(**D1**): "Fight for freedom or acquiesce to authoritarianism!"
(**D2**): "Do not acquiesce to authoritarianism!"
(**D3**): "Therefore, Fight for freedom!"

(D1)-(D3) appears to be a "valid" inference. However, it cannot be valid in the classical sense, according to which an argument is valid just in case the truth of its premises guarantees the truth of its conclusion. This is because both the premises and conclusion of (D1)-(D3)—being imperatives—are not truth-evaluable. The sense in which (D1)-(D3) is "valid" seems tied to the fact that the satisfaction of (D1) and (D2) guarantees the satisfaction of (D3), in a sense analogous to how the truth of (C1) and (C2) guarantees the truth of (C3). However, this would, at best, constitute a non-classical sense of validity.[14] And therein lays the point of contrast between a desire (conceived of as an imperative attitude) and believing, assuming, and fantasizing (conceived of as indicative attitudes). We do not need to resort to a non-classical notion of validity in order to make sense of the fact that the psychological content of believing, assuming, and fantasizing may feature in a valid inference. The apparatus of a classical truth-conditional logic is perfectly adequate for this task. It is the fact that the psychological content of believings, assumings, and fantasizings may all be captured by a truth-conditional logic that warrants their classification as theoretical attitudes.

To sum up, the two-content analysis offers us a conception of DOF that preserves the distinction between saying that an attitude has a certain DOF and saying that it has certain revision-conditions. According to the view presently on offer, an attitude is theoretical just in case it has indicative psychological content and practical just in case it has imperative psychological content. However, to say that an attitude has indicative psychological content is

---

[14] I include the qualification "at best" because it still remains controversial whether imperative inferences display their own kind of validity. This is the problem that Ross (1944) has dubbed as "Jørgensen's dilemma": how can imperatives (which are not truth-evaluable) feature in what appears to be valid inferences? Hence, we may characterize practical attitudes (i.e., attitudes with imperative psychological content) as those to which Jørgensen's dilemma applies. Theoretical attitudes, by contrast, do not face this problem, and are therefore not attitudes to which Jørgensen's dilemma applies. For an in-depth discussion of the semantics of imperatives, including arguments that an assertive interpretation of imperatives is inappropriate, see: Kaufmann (2012). Henceforth, I will leave the "at best" qualification implicit.

not yet to say anything about when that attitude is subject to revision. Hence, the two-content analysis allows us to say that believing, assuming, and fantasizing are all theoretical attitudes (owing to the fact that they may all feature in classically valid inferences) and also say that assuming and fantasizing do not aim at truth, and are therefore not subject to revision when they are false. On the present proposal, DOF is best conceived of not as a claim about when a certain class of attitudes is subject to revision, but as a claim about the types of inferences in which a certain class of attitudes may feature—indicative inferences in the case of theoretical attitudes, and imperative inferences in the case of practical attitudes.

References

Anscombe, Elizabeth. (2000) *Intention*. Cambridge: Harvard University Press.

Camp, Elisabeth. (2009) Two Varieties of Literary Imagination: Metaphor, Fiction, and Thought Experiments. *Midwest Studies in Philosophy: Poetry and Philosophy* **33**: 107-130.

Chan, Timothy (2013) *The Aim of Belief*. Oxford: Oxford University Press.

Engel, Pascal. (2004) Truth and the Aim of Belief. in D. Gillies (ed.), *Laws and Models in Science*. (pp. 77-97). London: King's College Publications.

Frege, Gottlob. (1879) Begriffschrift. in P. Geach and M. Black (eds.), *Translations from the Philosophical Writtings of Gottlob Frege* (pp. 1-20). New York: Philosophical Library.

Frege, Gottlob. (1979) *Posthumous Writings*. Chicago: University of Chicago.

Frost, Kim. (2014) On the Very Idea of Direction of Fit. *Philosophical Review* **123** (4): 429-484.

Humberstone, Lloyd. (1992) Direction of Fit. *Mind* **101** (401): 60-83.

Jørgensen, Jørgen. (1937) Imperatives and Logic. *Erkenntnis* **7** (1): 288-296.

Kaufmann, Magdalena. (2012) *Interpreting Imperatives*. New York: Springer.

Kissine, Mikhail. (2009) Illocutionary Forces and What is Said. *Mind and Language* **24** (1): 122-138.

Mastop, Rosja. (2011) Imperatives as Semantic Primitives. *Linguistics and Philosophy* **34**: 305-340.

Moran, Richard. (2004) Anscombe on 'Practical Knowledge.' in J. Hyman & H. Steward (eds.), *Royal Institute of Philosophy Supplement Volume 55: Agency and Action* (pp. 43-68). Cambridge: Cambridge University Press.

Owens, David. (2003) Does Belief Have and Aim? *Philosophical Studies* **115**: 283-305.

Platts, Mark. (1979) *Ways of Meaning*. London: Routledge and Kegan Paul.

Price, Huw. (1989) Defending Desire-as-Belief. *Mind* **98** (389): 119-127.

Ross, Alf. (1944) Imperatives and Logic. *Philosophy of Science* **11**(1): 30-46.

Smith, Michael. (1987) The Humean Theory of Motivation. *Mind* **96**: 36-61.

Searle, John. (1983) *Intentionality: An Essay in the Philosophy of Mind*. Cambridge: Cambridge University Press.

Searle, John and Daniel Vanderveken. (1985) *Foundations of Illocutionary Logic*. Cambridge: Cambridge University Press.

Stenius, Erik. (1967) Mood and Language-Game. *Synthese* **17**: 254-274.

Tuomela, Raimo. (2007) *The Philosophy of Sociality: The Shared Point of View*. Oxford: Oxford University Press.

Velleman, David. (1992) The Guise of the Good. *Noûs* **26** (1): 3-26.