

## Schmoozy Introduction

Peter Railton

University of Michigan

### *Introduction*

If *practical reason* is concerned with thoughtful normative regulation of action, then *theoretical reason* might be seen as a matter of thoughtful normative regulation of belief. The conclusion of a piece of practical reasoning, we are told, is an act or intention to act; the conclusion of a piece of theoretical reasoning, by parallel, would be a belief or a belief-tendency. Because theoretical reason is understood to be responsive specifically to *epistemic* – not merely pragmatic – reasons for belief, the norms involved in theoretical reason are norms pertinent to knowledge, including norms of evidence, justification, and truth. In both the practical and theoretical case we need not exactly agree with the conclusion another person reaches in order to attribute rationality to her – but it seems we do need to find *something* normatively appropriate about the means by which the other reaches her conclusion, given her capacities and circumstances. A somewhat fancier way of putting this: we need to see her as responding to the relevant sort of reasons in an apt way.

Moral philosophers have long found practical reason puzzling “by its very nature”. Practical rationality would appear to call for a kind of reasoning unknown to orthodox logic – a process Aristotle called “appetitive intellect or intellectual appetite”, of which the final step is the starting-point of action”. Appetite, as Aristotle would put it, belongs to the *alogon* part of the soul – but how then can we be intellectual about it, or deliberate with it, since intellect and deliberation belong to *logos*? Aristotle says that “Moral Science may be of great value to those who guide their desires and actions by principle (*kata logon*)” (NE 1095a16). But what would it be for desire, say,

to be guided by principle? How can *action tendencies* such as appetite, desire, or (for that matter) will, be subject to principles of valid inference?

At this point, we moral philosophers often look wistfully over to epistemology, where the phenomenon to be normatively regulated, belief, seems paradigmatically cognitive. There would seem to be nothing funny about “believing intellect or intellectual belief”, or about seeing in belief structures worthy of the name ‘valid inference’ or ‘evidence-based reasoning’. Doesn’t belief belong to *logos* – at least, those forms of belief that take propositional objects? A belief that *p* has the very same truth conditions as *p* itself. By contrast, a desire or intention to *q* is hardly made true by *q*’s truth. A belief that I am drinking water has the same truth conditions as *PR drinks water*. But it would be odd to assign to my desire to drink water the same truth conditions as *PR drinks water* – it often, perhaps usually, is the case that I desire to drink water in part because I am *not* doing so now. Suppose PR does get his drink. Is the desire *now* true? That’s odd to say, and drinking ordinarily makes a desire to drink go away, rather than gain validity.

To be sure, desires and intentions are often construed as having a propositional object: I seem to be liking in some motivationally-relevant sense the *idea* that *PR drinks water* – I act “under a description”. Moreover, desires and intentions are often criticized as “mistaken”, a seemingly cognitive and normative notion. But, by a time-honored tradition, this “mistakenness” is usually located not in the desiring attitude itself – which is seen as a non-cognitive state of liking or wanting, lacking truth conditions in any strict sense. Rather, it is located in the fact that the desire depends in part upon false belief. In Hume’s famous example, my desire or intention to eat a certain fruit might be said to be “mistaken” since I believe it to be tasty, when in fact it is not.

Yet it is just here that my problem arises. For even in the case of propositional belief, once we recognize that a belief is an *attitude* toward a proposition rather than a bare proposition, to make sense of a mistaken belief we need to make sense of the idea of a mistaken attitude. The mere fact that the object of the attitude is false does not suffice to show that the attitude is mistaken – as the the desire case shows, or the even simpler case of the attitude of denial. What is it about the belief attitude that makes it be the case that it is both necessary and sufficient for that attitude to be mistaken that its propositional object be false?

It is easy to confuse *mistakenness* with *defectiveness* of other sorts. For example, I might form my beliefs carelessly, or dogmatically, or with excessive credence. But they could still be *true*, even if defective epistemically from the standpoint of justification. To pick up on an earlier observation, they could also be free of this kind of epistemic defect – fully justified – and yet *false*. You could congratulate me on getting the facts right, or on being epistemically responsible, or, if I'm very lucky, on both.

In this sense, it won't do to say that we have an easy explanation of why the truth of the believed proposition is necessary and sufficient for its correctness from the fact that belief "aims at truth", and so a true belief is a *success* story. Belief does not aim at truth the way a practical agent aims at a goal. Practical agents can pursue a means to reach an end, but I cannot *believe* a proposition because I see it as an effective means toward an end. Perhaps believing that I have a useful idea about the nature of the belief attitude will motivate me to learn something about it, maybe even learn that my useful idea is false. But fully convincing myself of this doesn't induce belief that I do have a useful idea. If I'm to aim at truth as an end in the success sense, belief can't be my vehicle. Perhaps something weaker, like "acceptance" could, but the attitude of belief somehow isn't "fit" to play this role.

Sometimes people say that “aiming at belief” should be understood non-practically, as, say, “an attitude that represents itself as constitutively disciplined by norms of truth and evidence”. This sort of conceptually-laden, higher-order, self-reflexive representation seems like a lot to build into what it takes for my ten-year-old to believe that the floor of his room is solid, but set that aside for now. What worries me is a “regress problem” – it looks as if I will have to have beliefs to deploy this meta-representation within my mental economy. In order to “take” my attitude as one “disciplined by norms of truth and evidence”, I seem to need to believe certain things – about what my attitude is, how it is disciplined, what the norms of truth and evidence are, etc. So now I need to represent *those* beliefs to myself as “attitudes disciplined by norms of truth and evidence” in order to believe *them*. And on it goes. Belief had better “ground out” in something that isn’t itself a higher-order belief.

There is no *retro-version* of the regress problem, however. Once I have belief up and running, I can certainly come to believe that my beliefs are disciplined by norms of truth and evidence. Indeed, to the extent that I can do this reflective task successfully, I can gain greater and more robust confidence in my beliefs. We should be all for this sort of reflective self-scrutiny, and be sure to make it possible. The trouble is: it won’t be, if it is a *condition* for belief, i.e., if it is an account of *what it is to believe* in the first place. So we need to find a non-regress-inducing way for belief to get up and running, consistent with whatever truth there is in the slogan, “belief aims at truth”.

And I do think there is a substantial element of truth in it. Trying to figure this out *does* help us to distinguish belief from desire, and from such more voluntaristic states as acceptance or supposition. But notice that acceptance and supposition are also true or false iff the proposition accepted or hypothesized is. Yet neither “aims at truth” the way belief does. So, somehow,

we'll need to dissociate the explanation of how belief gets its correctness conditions from a requirement of *truth-aiming-ness*; but at the same time, we'll need to explain why this kind of *aiming-ness* doesn't "get in the way of" factual truth conditions, since aimingness does seem to get in the way of straightforward truth conditions in the case of desire. After all, in desire, aiming-ness allows the falsehood of the target proposition, say, *PR drinks water*, to be part of what explains why the desire is reasonable and appropriate – since I don't have a drink, and need one, I have every reason to want one. Nothing defective about *that* attitude, other things equal.

Begin with the simplest cases. When I simply "have a thought that *p*", there clearly is no aiming-at to get in the way of truth-evaluability, so that case seems straightforward. Once I start "supposing that *p*" or "hypothesizing that *p*" things begin to get more complicated. Now they're picking up a kind of steam in driving my mental economy forward, but still not aiming-at – I'm aiming-from, perhaps, primed to *start* inferring things from *p*, but under no pressure to do so. Since aiming-from is simply "taking to be true", we might think we see pretty straightforwardly how they supposition or hypothesis can only do their job if they add nothing to, and subtract nothing from, the simple truth condition. If not, then a critique could say, "Look, I asked you to suppose that *p* – but you've added something more, further suppositions. Stick with *p* and see what follows from it alone."

Now "belief that *p*" also involves "taking to be true", but it distinguishes itself from supposition and hypothesis precisely because it involves so much more commitment on my side – a host of attentional and inferential tendencies, expectations, action-tendencies, etc. are all brought to life. Now I *am* aiming. A belief that *p* is a trust in *p* that leads us to expect that things are the way *p* says, and count on this in action and inference. Yet somehow all this additional, attitudinal "whirl of organism", this feeling of confidence, bestirring of attention, arousing of

expectancy, setting up for a surprise if things don't turn out the way *p* says – *none of which is contained in or even hinted at in the content “that p” alone* – all this still adds or subtracts nothing that could block simple truth-conditions for belief. As Hume puts it, belief is a *firm conception* of a proposition that somehow “adds no new idea”, but nonetheless gives it an enormously powerful role in our lives. Though Hume sometimes speaks of this “firm conception” as involving a kind of vividness or liveliness, he recognizes that the phenomenology here is too thin and ineffable to help us understand belief's role in our lives. What matters is the functional role:

This variety of terms [‘firm’, ‘vivid’, ‘lively’, etc.], which may seem so unphilosophical, is intended only to express that act of the mind, which renders realities more present to us than fictions, causes them to weigh more in the thought, and gives them a superior influence on the passions and imagination. Provided we agree about the thing, ‘tis needless to dispute about the terms. [T 629]

How could belief be an attitude, a “feeling” as Hume puts it, fit for this exciting and activity-guiding role in our mental life, still preserving the straightforward truth conditions of the bare proposition? Wasn't it just this sort of “connection to activity” that led to non-cognitivist interpretations of moral terms, and to the original Frege-Geach problem? So why isn't there a Frege-Geach problem for belief? Whatever else we say about this belief attitude, it is not a *logical* entity, and truth-functional logic hardly applies to it (as is well-known). If practical reason yields *action*, and so cannot be theoretical reason, then theoretical reason, which yields *belief*, cannot be logic.

Philosophers of mind are probably blinking their eyes in astonishment at this point, or

rubbing away the tears of boredom or laughter. Here I am, rattling along as if this were the 18<sup>th</sup>-century, and we had no difficulty with naive talk about “ideas in the mind” – no “representation problem”, no “mental imagery debate”, no “problem of content”. True. I’m just assuming that I have the *object* of belief available, and that it has content and truth conditions. I’m working the other side of the street from Harty Field when he observed in his seminal essay on “Mental Representation” that belief is a mental state attitude compounded of at least two elements:

**attitude of believing [ mental representation of a proposition ].**

He gives the convenient label ‘believe\*’ to the attitude. So we might represent a belief that  $p$  as:

**Bel\* [  $\phi_p$  ]**

where  $\phi_p$  stands for a mental representation of the proposition  $p$ .

Field focuses primarily on the need to explain how an actual, embodied mind could be in the position of successfully representing such a thing as a proposition. Of the attitude believe\* itself, he writes:

The rough idea of how to give an account of [“what it is for a person to believe\* a sentence of his or her own language”] should be clear enough: I believe\* a sentence of my language if and only if I am disposed to employ that sentence in a certain way in reasoning, deliberating, and so on.

We seem to know pretty well, then, the “job description” for believe\*, and what else is there to know? Why not say that there is a “belief box” somewhere in the mind’s functional architecture, and once we can figure out how to get a contentful representation of  $p$ ,  $\phi_p$ , available to the mind to put into it, we’re home and dry? The psychologists will tell us how a belief box gets built out of the human sub-psychology; the philosophers will tell us how a contentful representation with truth conditions could be available to put there.

What is *my* problem about belief? My problem is that I don't see that we philosophers have done our job yet when we talk about "belief boxes" and the like. We need to say how a "belief box" could work, in terms of a kind of universal sub-psychology of belief. Imagine trying to solve the problem of moral judgment by saying that there is a "moral judgment box", and once we have a semantically meaningful representation of an act, we can judge that act right or wrong simply by plunking it in that box – leaving the sub-personal plumbing to empirical psychologists. What about the quasi-*a priori* truisms associated with moral judgment? For example, that there is some sort of link between making a moral judgment and being motivated? If a propositional content becomes a moral judgment by being plunked into the "moral judgment box" then why can't moral judgments retain the proposition's straightforwardly cognitive truth conditions? They did in the "belief box", after all, and it was hooked up in all kinds of ways with the organism's activity.

Perhaps they can retain the truth conditions. But then how can there also be a semantic or quasi-analytic or *a priori* necessary tie to *motivation*? It looks as if being in the "moral judgment box" does add some content to, or subtract some factuality from, the representation; so orthodox truth conditions become problematic. The history of 20<sup>th</sup>-century meta-ethics follows: Moore's "open-question argument" and intuitionist Platonism; dissatisfaction with which led to Ayer's and Stevenson's non-cognitivism; dissatisfaction with which led to Hare's prescriptivism; dissatisfaction with which led to Gibbard's norm-expressivism and Blackburn's projectivist quasi-realism, on the one hand, and McDowell's and Wiggins' "sensitivity theories" on the other, not to mention countless "naturalistic" projects.

Why all the dissatisfaction? One reason, certainly, none of these accounts seemed to capture the simple *objective purport* of moral judgment – it seems to claim not only a kind of

objective validity, but the kind that is truth. The “Frege-Geach problem” is a manifestation of this – moral claims appear to belong in their own right in all the roles of *bona fide* propositions. So are issues about supervenience, understood as Hare or Gibbard did: moral assessment *must* be answerable to the non-moral facts. So are issues about moral disagreement, moral epistemology (we seem to have no special “moral sense” or “moral epistemology”). In short, moral judgments *did* behave just like beliefs, i.e., behaved as if they had straightforward truth conditions, permitting all the usual logical operations, justification and confirmation moves, etc. To be sure, there *also* seemed to be something subjective there, that could not be “nailed down” by ordinary factual inquiry. A kind of openness to contestation that was substantive, unlike Moore’s conceptual version of the “open question test”. And moral judgment did seem to have some quasi-analytic or *a priori* necessary tie to motivation. So there seemed also to be a sense in which straightforward truth conditions looked “problematic in principle” – a fact/value distinction, an *is/ought* distinction.

For facts, *is*’s, and beliefs, by contrast, had no such tie to motivation. I can believe that *p* while being in favor of this, or wanting to see it continue; I can equally believe that *p* while being opposed, and deeply committed to changing things. Either way, the truth or falsity of *p* – the “factuality” or “*is*-ness” of *p* – settles the correctness of my belief attitude definitively. Belief looks motivationally *inert*, but moral judgment is not like this. Hence the ethicists’ close philosophical scrutiny of what it would take to build a “moral judgment box” such that a propositional representation could be put into it, and box-and-content together could play the *action-guiding role* attributed to moral judgment. But no one has yet built such a box, most would agree. The best account of action-guidingness – norm-expressivism – leaves us with the worries about objectivity and with the Frege-Geach problem still to solve. The best account of

how to solve those problems – cognitivism – leaves us with the “judgment internalism” puzzle unsolved. For how can motivation be “built into” a purely cognitive judgment, like a belief? Perhaps it is an ordinary belief with an extraordinary *content*? – Maybe an “action-guiding” or “essentially normative” *concept* is in play. But now we’re back to facing the problem of a statement containing such a fancy concept could have purely *factive* truth conditions while still working its magic on motivation.<sup>1</sup> Even supervenience begins to come into doubt, and we need to start populating reality with “normative facts” to answer to these special concepts.

Briefly, my position will be that this trip is not really necessary. The action-guiding force has to come “from our side”, not from the side of the concept or from some independent domain of action-guiding facts. To think a concept or a fact could have this power is akin to commodity fetishism – thinking that the exchange value of a good, its power to obtain other goods in exchange and bring other goods (like labor) into play, somehow resides in *it*. Hume and Locke argued that the power of a currency to command goods and services resides in our *acceptance* of it, not in the currency itself. Marx and Sraffa argued that the power of commodities to call into being labor or serve as capital resides in a *social system*, not the commodity itself. Wittgenstein argued, crediting Sraffa, that the power of words or logic to regulate our practices resides in our *dispositions to use them in social practices of deference*, not in the formal apparatuses alone; Fodor argued that the power of check to call off the creditors resides in a *banking system*, not the scrap of paper. All these people strike me as right. If there is any *rule-following* going on, then that is because there is *following* going on, and the rule does not supply that, we do.

---

<sup>1</sup>Some philosophers have gone so far as to suggest that only certain very special people – virtuous ones – can really grasp these special concepts, so *of course* moral motivation is guaranteed by factual application of the concept. The rest of us somehow can’t get it – like sightless individuals who can’t quite get the concept of ‘redness’. One problem is that moral notions don’t seem to have a peculiar phenomenology, like colors, that could guide judgment. But then, maybe *I* can’t see that phenomenology, as a sightless person can’t see colors?

This seems to throw us back to norm-acceptance. There is norm (rule), there is acceptance (following). So moral claims, for example, *express* acceptance of norms. Trouble is, how can a non-propositional expression of a psychological state of norm-acceptance have plain vanilla truth conditions, just like fact-stating discourse? Expressivists emphasize – repeatedly, since they are often misunderstood on this point – that they see moral judgments as *expressing* attitudes, not as *reporting* them. Reports of my psychological state might have plain vanilla truth conditions, but then that isn't what would do the job of guiding my action – since I can coherently think I *ought not* to follow my psychological state, *ought not* to simply follow my feelings.

This distinction seems necessary, too, if we are to account for moral disagreement. I am in the psychological state of accepting norm *N*; you are in the psychological state of accepting norm *never N*. We sound off, reporting our condition. No disagreement there. But I think we *ought to defer to N*, and you think we *ought not*. So we disagree alright – our attitudes share a common subject matter – *what to do* –, and thus compete in a way that our reports do not.

We are left hunting for a solution to Frege-Geach worries.

But it is much worse. For most of the 20<sup>th</sup> century, it seemed as if this funny business about *normative judgment* was the headache of ethics and aesthetics – “value theory”. Lucky for the rest of philosophy, they had honest work, like epistemology and semantics to give us some factual modes of discourse to work with, and pedestrian truth values to light our path. Ethicists assumed this, too. Anything they could “solve” using semantics, say, would allow moral philosophy to be a cognitive exercise – they could discover facts of logic and meaning even if they couldn't *in principle* discover moral facts. But closer reading of the situation undid this comforting thought. Epistemology discovered that “justification” and “evidence” are normative

concepts, that “belief” is a “constitutively-normative, inference-guiding attitude” rather than an “inert representation”, and that “rational inference” is not logical deduction but a norm-guided process of reasoning. ‘Kripkenstein, with the help of Bogopke, discovered that “meaning” and “content” are normative. We all seem to be in the same boat. So if Frege-Geach, for example, is a problem, we’ve got it in spades. Even in “theoretical reason”.

But maybe we got off on the wrong foot because of some seductive ways of thinking about the problem of action-guidingness or normativity, arising from the early focus on the special case of ethics. And maybe the core problem about normative guidance has been in philosophers’ sights for years and years – even Aristotle, in his critique of Plato, might have seen it. And maybe these giants – Aristotle, Hume, Kant, Wittgenstein – have shown us a way of tackling it in part by helping us with just the problems I’ve been trying formulate – the nature of a “belief box”, a “moral judgment box”, a “desire box”, or an “intention box”. That preposterous claim is the one I’ll be defending. This paper is an effort to begin to show how. It works by rehearsing various things to set up the problem, then making a shameless appeal to historical authorities, and then trying to state a principled position. Along the way, I lean heavily on Hume’s view that belief is a “feeling” that plays a powerful role in our lives. That is, that belief is not an “inert representation”, but has a kind of ert-ness that *requires* it to have the truth-conditions of its propositional object. I also lean heavily on Wittgenstein’s notion of “following a rule blindly”. The latter I had tried to say a bit more about in the second paper I’m enclosing, on *a priori* rules. The former – the bit about belief – I have defended in various papers, but I won’t dump any more Railton-prose on you than the “Moral Realism, Moral Belief” paper and the paper on rules.

Looking forward to your well-known searching scrutiny – as you’ll see in these papers,

I'm a big believer in feedback mechanisms as essential to normative self-regulation. Please help me to practice what I preach.