

Gibbard, Allan. *Meaning and Normativity*.

New York: Oxford University Press, 2012. Pp. 336. \$55.00 (cloth).

forthcoming in *Ethics*

In this book, Allan Gibbard develops a ‘normativist’ and expressivist account of meaning, according to which meaning ascriptions concern how we *ought* to use language, and are explained as complex *plans* for using language in certain ways. The book, which is in significant part based on articles that Gibbard has published on the topic over the past two decades, or so, is the clearest and most thorough examination of this combination of ideas to date. Gibbard does not much discuss other normativist views or the criticisms of normativism, but rather focuses on outlining his own distinctive approach. A part of the project is to use expressivist metanormative theory to offer insights into issues in philosophy of language; another part is to continue developing Gibbard’s influential expressivist metanormative theory through examining the implications for expressivism of the idea that the concept of meaning would be normative.

The book is rich in original ideas and arguments, and the topics canvassed or commented upon are significant and bewildering in their number: Gibbard offers, for instance, normativist-expressivist explanations of the concepts of synonymy, analytic and *a priori* equivalence, and reference; a solution to Kripke’s Wittgenstein’s ‘skeptical paradox’; commentary on Paul Horwich’s use-theoretic view; an explanation for objective oughts in terms of subjective ones; updates on his expressivist view; an articulation of the difference between expressivism and non-naturalist realism in metanormative theory; and a response to Mark Schroeder’s discussion of the Frege-Geach problem. It is a compact and difficult book, but serious students of the relevant topics should find its study rewarding, and clearly it is essential reading for anyone working on meaning and normativity.

Gibbard distinguishes between the ‘weak’ and the ‘strong’ theses of the normativity of meaning. The weak thesis says that *means* implies *ought*. For example, if I mean what you mean by ‘Snow is white’ and ‘Nothing is white,’ I ought not to accept both of these sentences at the same time. The strong thesis says that “meaning is fully definable in normative and naturalistic terms.” Gibbard follows A. C. Ewing in taking the concept of ought as the primitive normative concept (10–15, 53–55).

Although Gibbard’s focus is not on arguing for normativism, but rather on how a normativist account of meaning should perhaps be developed, he does give some “preliminary support” to the idea in chapter 1. One reason for accepting the weak normativity thesis is that it

would solve the “automaticity problem” (150), that is, “explain why certain oughts follow from claims of meaning invariably” (13). These oughts are not just a matter of it being desirable to believe the truth. An evil demon could see to it that believing both that snow is white and that nothing is would be for the best, but there is a sense in which it would remain true that one ought not to have both of these beliefs. This, Gibbard suggests, follows from the meaning of ‘nothing’ (13–16).

Gibbard acknowledges that his discussion of the automaticity problem leaves room for alternative solutions. Some argue that it follows from the concept of a *belief* that one ought to believe that p only in case it is true that p . Perhaps this would help to explain why it is that one ought not to believe both that snow is white and that nothing is. However, another important motivation for accepting the normativist view, also in its stronger version, is provided by the idea that attempts at explaining meaning non-normatively – *via* a purely naturalistic reduction, say – have proved problematic (16–17).

A third reason for accepting normativism comes up later in the book. Some alternative views leave room for considerable indeterminacy of meaning. Our actual, “messy” dispositions to use language in certain ways, for instance, may fit many idealized patterns of use (a use-theoretic candidate for explaining meaning) equally well, but the normativist view helps to remove some slack here (106–110, 150).

In chapters 2 and 3, Gibbard first suggests that it is important to distinguish between the *concept* of meaning and the meaning *property*, the former of which, and not the latter, is normative. He then proceeds to argue that once this is kept clearly in mind, space opens for ‘straight’ solutions to Kripke’s Wittgenstein’s (KW) paradox. KW is often taken to attack naturalistic accounts of meaning properties on the grounds that meaning is normative. Someone might suggest that I mean addition by ‘+’ in virtue of having certain dispositions, but KW’s ‘meaning skeptic’ may agree that I have these dispositions, and yet think that I mean rather quaddition (a function which gives the sum for numbers under 50, and the answer 5 for bigger numbers). KW seemingly concludes that facts about my dispositions fail to “tell me what I ought to do,” and so fail as candidates for the meaning property. (Saul Kripke, *Wittgenstein on Rules and Private Language* [Oxford: Blackwell, 1982], 11–24.)

Gibbard points out that this is to draw hasty conclusions about meaning properties on the basis of conceptual considerations (cf. the standard rendition of Moore’s Open Question Argument). Suppose that I am only learning to use ‘+’, and that, when asked questions of the form ‘ $x + y = ?$ ’, I reliably respond with the sum, except for numbers larger than 50, for which I respond randomly. Suppose, moreover, that I live in a community of quadders. Someone – call

him Jerry – might think that given my dispositions (and certain simplicity considerations), I mean PLUS by ‘+’. Another theorist – call him Tyler – might think that given my community’s dispositions, I mean QUUS. What is at issue between Jerry and Tyler? Perhaps this: whereas Jerry thinks that I ought, when asked ‘ $57 + 68 = ?$ ’, to respond ‘125,’ Tyler thinks that I ought to respond ‘5’. Jerry and Tyler may agree about the dispositional facts of the case, and yet disagree on which naturalistic properties ground the relevant oughts – on which naturalistic properties constitute the meaning properties. One of them could be getting the relevant oughts (roughly) right. There is no quick way of rejecting naturalistic proposals by appealing to the meaning skeptic.

There is of course much more to KW’s discussion of the naturalistic views than this, and Gibbard discusses some of KW’s other arguments in some detail (chapter 3).

An important corollary of Gibbard’s discussion (chapter 2) is that a clear distinction should be drawn between normativity and *sociality* theses. Jerry accepts that meaning is normative, but denies that it is essentially social, and one could accept a view like Tyler’s minus the normativist claim.

In chapter 4, Gibbard argues that the oughts constitutive of meaning must be understood as subjective, as “a matter of what one ought to do in light of one’s evidence” (21). From snow’s being white it analytically follows that one ought, in the objective sense, to believe that snow is white. This is puzzling, for now it seems like snow’s being white would be a normative matter, which it clearly is not. Gibbard suggests that the relevant ought-claim “is a normative claim only degenerately,” and proposes to explain it, and objective oughts more generally, in terms of subjective oughts.

The claim that A ought, in the objective sense, to φ , is analyzed as meaning that A^+ ought, in the subjective sense, to φ – where A^+ is someone who ought, in the subjective sense, to self-attribute all that is the case with A , as A in fact is. From this analysis it follows immediately that A ought, in the objective sense, to believe that p if and only if p . The analysis may seem dubious. Suppose that A has no beliefs as to the strength of the enemy, but that in fact the enemy has assembled five depleted regiments. A^+ ought_{subj} to believe that the enemy has assembled five depleted regiments, and that she has no beliefs as to the strength of the enemy. This seems questionable, as these two things cannot be jointly truly believed (cf. Krister Bykvist & Anandi Hattiangadi, “Does Thought Imply Ought?” *Analysis* 67 [2007]: 277–285). Gibbard, however, suggests that this is unproblematic (see also his online supplement to the book, “Correct Beliefs as to What One Believes: A Note,” <http://www-personal.umich.edu/~gibbard/Bykvist-Hattiangadi2012jl-pap-bel-truth.pdf>).

Gibbard also thinks that, given that A^+ could safely go and observe the strength of the enemy, this is what she ought to do (she believes that she does not know the strength of the enemy, after all). But it is not obvious that we get this result. Would it not be equally sensible to say that A^+ ought to go quickly report the strength of the enemy (she believes that the enemy has assembled five depleted regiments, after all)? It seems unlikely that A^+ ought to do both of these incompatible things.

In chapters 5 and 6, Gibbard proceeds to examine Horwich's account of meaning, and to outline his own normativist theory, which gives a normativist twist to some of Horwich's ideas. Many terms are treated as theory-laden. Let $T(\text{phlogiston})$ be the phlogiston theory. The *Carnap pair* for this theory says that (i) if there is anything that plays the role that the theory says phlogiston plays, then that is phlogiston, and that (ii) if nothing plays this role, then there is no such thing as phlogiston. On Horwich's view, very roughly, my meaning PHLOGISTON by 'phlogiston' is a matter of my implicitly accepting this Carnap pair. Gibbard's suggestion is that it is, rather, a matter of it being the case that I ought to accept it. Moreover, it must be the case that I ought to accept it *a priori*, regardless of my evidence, and as *analytic* – that is, “on any intelligible supposition” (116–117, 123–124).

In giving the meaning of 'phlogiston,' some other terms are taken as understood. The meanings of all these terms are, then, given by sequences of Carnap pairs. Two words – my 'bachelor' and Pierre's 'célébataire,' for example – are synonymous if their meanings are fixed by matching sequences, and synonymy for sentences requires both synonymous elements and “matched syntactic structures” (119–128).

A tough question is: how do the sequences of Carnap pairs get started? Gibbard offers some sketches of possible explanations for the meanings of the more basic items (e.g. terms directly tied to experience), but admits that his proposals will require far more development than is given in the book (128–132).

The normativist account of *reference*, proposed by Gibbard, says, very roughly, that Ada's use of 'I' in her claim 'I am sad' refers to Ada just when a certain kind of pattern of oughts obtains – a pattern instantiated, for example, by this: if Ada ought to accept 'I am sad,' I ought to accept 'Ada is sad.' This idea is qualified and defended in detail in chapter 7.

Having outlined a theory of meaning in terms of a primitive ought, Gibbard proceeds, in chapter 8, to give an expressivist account of this concept. On the resulting view, to think that Pierre's 'Les chiens aboient' means the same as my 'Dogs bark' is to plan to form beliefs with the former sentence, if in Pierre's shoes, the same way that one would plan to form beliefs using the latter, if in mine.

“Plans,” here, cannot be just ordinary plans. We do not form beliefs at will, and we cannot plan for the kinds of fantastic circumstances we know our plans could not affect (the case of being in Pierre’s exact subjective circumstances, say). Gibbard suggests that his “plans” are *preferences* for hypothetical cases (176). This solves the latter problem, but how can we “plan” our beliefs? Gibbard proposes that a “plan” to believe that p is to a belief that p as an ordinary plan to perform φ is to willing to φ . But it is not quite clear how this is supposed to work. A plan for action (setting off at noon) may lead to willing some particular act (e.g. taking bus 67) by rationally causing, or constituting, this willing in combination with an instrumental belief. But Gibbard probably would not want to say this about planning to believe and believing. It is not quite clear, then, how “plans” should be understood. They answer the question of what to believe, alright – but that sounds a bit like saying that they are judgments concerning what we ought to believe (which is what Gibbard is trying to explain).

Gibbard then turns to asking what follows for expressivism, if what has been said in the earlier chapters is on the right track. Chapter 9 tests Gibbard’s metatheory of meaning in its own terms. Gibbard argues convincingly that the theory’s being normative ‘all the way down’ is no problem, and offers normativist-expressivist accounts of the Moorean idea that normative terms cannot be analyzed in terms of naturalistic ones, as well as of expressivism itself. This last idea has raised some concerns in the philosophical community, but Gibbard nicely explains how it might work. Accepting expressivism might amount, first, to holding that a sentence of the form ‘ A ought to φ ’ is analytically equivalent to the conditional imperative ‘If I am A , then do φ !’, and, second, to accepting that we can give a certain kind of naturalistic story, perhaps in terms of dispositions to act, of the property that meaning OUGHT consists in.

Gibbard discusses at some length the challenge posed by *akrasia* (207–217). On his view, if I accept that I ought, right now, to call my doctor, I plan to call my doctor. Many would object that I may accept the belief without accepting the plan. In his earlier work, Gibbard bites the bullet, but he now wants to make room for this phenomenon. First, he points out that his account allows for the possibility of thinking ‘I ought to call now,’ and yet failing to *call*. This does not, however, address the worry that it is possible for me to think that I ought to call, and yet fail to *plan* to do so. Gibbard replies that we should distinguish between normative and ‘simulational’ concepts of meaning, where the latter (unlike the former) allow for the possibility of my thinking that I ought to call my doctor while not planning to do so. Not much is said to clarify this, but Gibbard is explicit, of course, that the proposal needs further development (210–212, 216–217).

In chapter 10, Gibbard examines the relationship between expressivism and non-naturalist realism. Given enough quasi-realist manoeuvring, expressivists may accept that we

have normative beliefs, some of which get the normative facts right (230, 232–233). Someone like Scanlon, on the other hand, may agree with the expressivist-sounding claims that ought-judgments normally motivate, and that *ought* entails *do!* (230–232). A potential remaining difference would then be that non-naturalists wish to explain normative judgment as a discernment of how the world is, normatively speaking. They have yet to adequately explain, however, why it is that pursuing reflective equilibrium in normative enquiry tends to lead us, as products of natural selection, to views that would track the normative truth (236). Expressivists, who do not explain normative judgments as representations of the normative facts, need not shoulder similar explanatory burden (236–240).

Gibbard closes the book with two substantial appendices. In Appendix 1, he argues that objects of belief are not propositions, and asks how we should understand communication, given that this is so. Appendix 2 engages with Schroeder’s discussion of the Frege-Geach problem. The problem is roughly that of explaining what is expressed by arbitrarily complex normative sentences. In his earlier work, Gibbard seems to suggest that we understand the states expressed in terms of what states of mind they would *disagree* with. As no explanation is given on the disagreement relation, this seems unsatisfactory. The characterizations of the relevant states in terms of disagreement relations may be correct, but do not give adequate, specifically *expressivist* explanations of them. Moreover, on Schroeder’s view no satisfactory expressivist account along these lines is going to work, as it would need to appeal to mysterious ‘*B*-type’ inconsistencies, not just to ‘*A*-type’ inconsistencies that hold between two attitudes of the same type with inconsistent contents, and that everyone needs to account for. Finally, in order to make sense of all the different complex constructions, Gibbard would need to make sense of an infinite hierarchy of *B*-type inconsistencies. (See e.g. Mark Schroeder, *Being For* [Oxford: Oxford University Press, 2008].)

I am far from confident that I have understood Appendix 2 correctly, but Gibbard now seems clear that disagreement is to be understood as a *state of mind*, and not as a relation holding between such states (274–275) (cf. Ralph Wedgwood, “Schroeder on Expressivism: For – or Against?” *Analysis* 70 [2010]: 117–129). He then assumes that disagreeing with a state (action) and being in this very state (performing this action) is inconsistent. This, he suggests, is an okay assumption; something – be it disagreement or, say, truth – will be left unexplained by everyone (277–282). This is *B*-type inconsistency, but that is fine: everyone needs to explain *some B*-type inconsistencies, like that between preferring *A* to *B* and being indifferent between these options (284–286). Suppose, now, that we may take disagreement and ‘combining’ states of mind as understood. We can then explain the truth-functional constructions: to think that one ought to

drink is to disagree with not drinking; to think that it is not the case that one ought to drink is to disagree with disagreeing with not drinking; to think ‘one ought to drink or one ought to smoke’ is to disagree with combining (disagreeing with disagreeing with not drinking) with (disagreeing with disagreeing with not smoking). And so on (274–277, 289–290). This story would not result in an infinite hierarchy of *B*-type inconsistent attitudes, and the assumptions made seem relatively modest.

If this is the right way of reading Appendix 2, however, what Gibbard is suggesting is a version of a ‘higher-order attitude approach’ to the Frege-Geach problem, and may inherit its familiar problems. For instance, thinking something of the form $p \vee q$ will, on Gibbard’s view, turn out to be equivalent to thinking something of the form ‘One ought not to combine thinking that not p with thinking that not q ,’ and this seems problematic (but see Schroeder, “Higher-order Attitudes, Frege’s Abyss, and the Truth in Propositions,” in Robert Johnson & Michael Smith (eds.): *Passions and Projections* [Oxford University Press, forthcoming]).

There is more to Appendix 2 – some comments on quantification and propositional attitudes other than belief, for instance. Like every chapter preceding it, it contains a lot that will hopefully be subject to much discussion in the years to come.

TEEMU TOPPINEN
University of Helsinki