

10/17/2006

Jackson—Week 6: “Looking for Clues at the Scene of the Crime”

1. Plan: Do the session on two big parts (before and after the break):
  - a) Before: The argument of Chapter 5, descriptive vocabulary and describing. Macarizing and “disciplined syntacticism”.
  - b) After: “moral functionalism” (what *theory* do we Ramsify? Should we be K-functionalists in his sense for the meaning of *every* vocabulary? Centralism and the issue of the *semantic autonomy* of the vocabularies); the possibility of natural-kind skepticism about the moral; what we ideally would desire theory of the descriptive content of moral vocabulary. Here there are two main foci:
    - i) Moral functionalism in general.
    - ii) FJ’s particular proposal for an *a priori* descriptivist analysis of moral concepts.
2. The ambition of these last two chapters: to solve the location problem for ethics, to show that in making ethical claims we are describing the world. (One of the biggest issues to which 20<sup>th</sup> century ethical theory devoted itself.)
3. From *scientific* naturalism to *descriptive* naturalism:
  - a) Q: What is the difference between the issue of the supervenience (and hence, he claims, the analytic definability) of the *moral* in Chs 5 &6, on the one hand, and the *psychological* in Ch2? (This is a difference shown, for instance, in the *very* different structure of the arguments he employs. He does not, for instance in the later argument appeal to his “*minimal duplicates*” version of supervenience (about which I complained in connection with his Chapter 2).
  - b) Notice that here it is *definability* and *derivability* (entailment) of E-facts from D-facts that FJ wants—hence a version of full *reducibility*.
  - c) A: He has moved from *scientific naturalism*, in the form of *physicalism*, with respect to *psychology* to *descriptive naturalism*, in the form of an *ideal conditions humean internalism* about the *moral*. (Jackson says in Ch. 6 that he will call this view “*descriptivism*”, by contrast to “*naturalism*” so as not to risk any confusion with scientific or physicalistic naturalism about the psychological.) But he says things like [146]: “For it is common ground with the *ontological descriptivists* that there is no *ethical nature* over and above *descriptive nature*.” (He there contrasts his “*analytical descriptivism*” with *ontological descriptivism*, as two species of “*metaphysical descriptivism*.”) [147]: **“Metaphysical descriptivists think that how the world is, how we take things to be, and conventions of word usage, can be exhaustively given in descriptive terms.”** (Cf. (13b) below.)
  - d) Q: What is the connection between these, that is, between the *scientific* and the *descriptive* forms of naturalism?
  - e) A: It is provided precisely by the *scientia mensura*: Sellars’s principle that “in the dimension of *describing* and *explaining*, *science* is the measure of all things, of those that are, that they are, and of those that are not, that they are not.” (It is to the essay that enunciates and expounds that doctrine, Sellars’s “*Philosophy and the Scientific Image of Man*” that we turn our attention next week. That fact is enough to make one believe in *pedagogic providentialism*: the idea that our readings are being guided by

some wise and benevolent intelligence working out a plan set out before the beginning, seeing to it that there is some “one, far-off divine event, towards which the whole creation moves.”—or, less grandly, just trying to see to it that everybody has a reason to “tune in next week at the same time.”)

- f) Q: Why does FJ think there is such a big difference between showing the supervenience of the psychological on the physical (Chapter Two) and the supervenience of the moral-normative on the descriptive (Chapter Five)?
- g) It is *not*, I think, because moral facts or values, thought of as part of the furniture of the world, would be, in Mackie’s phrase, “queer” [a term that suggests the advent of a new sub-discipline with queer studies: its analytic metaphysics wing]. What would be ‘queer’ would be if such things were causally isolated from everything else, or otherwise *ectoplasmic*. But this does not seem to be what motivates FJ’s procedures here.
- h) A: Because in his set-up, (serious) metaphysics is about showing how one set of *facts* makes another set of *facts* true. The location problem takes place entirely between *descriptive* vocabularies: ones that try to say *how the world is*, that is, what is *true*, what the *facts* are. If moral vocabulary is not in that line of work, then there is no location problem for it, as there is not for, say, *commands*. For there is no question of what makes them *true*: nothing does, none of them are. We can only ask when that speech act is *appropriate*—and that is not a question of (serious) metaphysics.
- i) So he addresses moral vocabulary on the assumption (he calls it “cognitivist”) that moral vocabulary *does* state facts, *is* truth-evaluable, and hence raises a location problem. One question we must address is: what is the relation between facts, description, descriptive vocabulary, declarative sentences, and truth-evaluability?
- j) That is why he can, in effect, take supervenience of moral vocabulary on ‘descriptive’ vocabulary for granted, as a starting-point. For if there is not even supervenience on the descriptive, then there is no location problem. (But notice that he *does* claim [147] that “conventions of word usage” can be fully specified in descriptive terms, so he probably thinks that what one is *doing* in offering, say, commands, *is* something fully describable.)

#### 4. I: The argument

- a) Here is the Master Argument of Chapter 5, which takes us from supervenience to definability (notice that FJ says that Simon Blackburn takes this argument to be a *reductio* of cognitivism, arguing by *modus tollens*):

[122-3] E is a sentence framed in descriptive and ethical vocabulary. “Each world at which E is true will have some *descriptive nature*...And for each such world, there will be a sentence containing only descriptive terms that gives that nature in full. Now let  $w_1, w_2, \dots$  be the worlds where E is true, and let  $D_1, D_2, \dots$  be purely descriptive sentences true at  $w_1, w_2, \dots$ , respectively, which give the full descriptive nature of  $w_1, w_2, \dots$ . Then the disjunction of  $D_1, D_2, \dots$ , will also be a purely descriptive sentence, call it D. But then E entails and is entailed by D. For every world where E is true is a world where one or the other of the  $D_i$  is true, so E entails D. Moreover, every world where one or the other of the  $D_i$  are true is a world where E is true, as otherwise we would have a violation of (S) [the supervenience of the ethical on the descriptive]. Therefore D entails E. The same line of thought can

be applied *mutatis mutandis* to ethical and descriptive predicates and open sentences: for any ethical predicate there is a purely descriptive one that is necessarily co-extensive with it. **It follows that ethical properties are descriptive properties.”**

b) FJ does not need a *closure* condition here *within* the definition (as he did with his notion of *minimal physical duplicates*). It is enough if we stipulate that we are talking

c) The big issue is: how many possible worlds are there? Not, presumably, a *finite* number. That is why FJ allows elsewhere that the disjunction may require an infinitary language. Nor does he assume that the members of this infinite conjunction can be *recursively* specified, or are even *recursively enumerable*. (If, as I'll argue below, *mathematical* vocabulary is ‘descriptive’, then we have some very large cardinals indeed to specify already in the *conjunctions* that express the “descriptive nature” of the worlds.) But are there only a *countably* infinite number of possible worlds? [Notice that he has built a countability assumption *into his subscripts*.] If one is a physicalist, this is an empirical question. If Newtonian worlds are *possible*, then particles can take *continuum* many positions, momenta, and so on. And then there is an *uncountable* infinity of possible worlds. Then the disjunction has to be *uncountably infinitary*. But what do languages like this even look like? How do quantifiers work in them? Are they still “first order”? What do we use as the *expressions* in an uncountably infinite world?

d) But there are some really important philosophical issues raised by his discussion. And we can worry about them without having to hold our consideration hostage to his Master Argument.

5. I: Describing

a) What does describing contrast with? Evaluating? What about explaining? Giving reasons? Supposing? Emoting about?... Is there any definite totality of such things one can do with language? And don't their vocabularies overlap substantially? Think of Wittgenstein here: how many speech-act kinds that contrast with *describing* are there? Is it some definite totality?

b) Is there such a thing as “descriptive vocabulary”? Is there such a thing as “hypothetical” or “suppositional” vocabulary? Is there a vocabulary of recommending, or ordering, or explaining? In any case, there is a substantive move from distinguishing the *activity* (speech act kind) of *describing* to the notion of a distinctive vocabulary of description: descriptive vocabulary. Compare *observing* and *observational* vocabulary. There is no vocabulary *all* of whose uses are observational, in the sense of making non-inferential observation reports. Everything that can be used to do that can also be used to state the conclusion of an inference. (Inferring is a distinctive discursive activity. Is there such a thing as *inferential* vocabulary? In a sense, yes: *theoretical* vocabulary, which can *only* be applied as the conclusion of an inference.) By “observational vocabulary” we mean vocabulary that has *any* observational use. It contrasts with *theoretical* vocabulary, which does not. (But cf. Sellars's point that this boundary is porous and fluid, subject to change over time.) There is also no vocabulary *all* of whose uses are *deictic* (demonstrative), and very little vocabulary that *cannot* be used deictically.

The point is that the *move* from the *activity* of *describing* to the definition of a distinctive kind of *vocabulary*, *descriptive* vocabulary, requires both a substantive definition of the relationship between them, and some real justification.

c) The issue is whether there is a distinctive range of *facts* that are *descriptive* facts. Does every statement, in *stating* something, purport to *describe* how things are? Consider *mathematical* vocabulary. It can be used to describe *structures* of things. So the three-dimensionality of space can be described by an equation of rotations:  $ij = k$ . Am I describing when I say that the knife is sharp? Surely. That if you cut a roast this nub here on the handle rubs uncomfortably on your hand? Yes. What about expensive (is that a description or an evaluation)? That the handle is badly designed? That the knife is well or badly made? That it is a *good* knife? But what if the apparently *evaluative* (but still, notice, *declarative*) statement that it is a *good* knife is just *equivalent*—even, *analytically* equivalent to some conjunction such as: “The knife cuts smoothly, effectively, and efficiently, keeps its edge, is easily sharpened, fits comfortably into the hand, and is easy to use.”?

d) Consider what we would need to do if we were not by law allowed to *evaluate* a candidate, but could only describe her. Can we call the candidate *honest*? (FJ sees this as a borderline case.) Can we say that she is *sincere, reliable*, that she usually does what she says she will do? Can we describe her as a *junkie*? A *whore*? A *Republican*?

e) The implicit model—made explicit by some philosophers, often those who think of themselves as following Hume—is that proper inferences can only go from descriptive premises to descriptive conclusions, and that if an evaluative conclusion is to be drawn, evaluative premises are required. Davidson is a good example.

f) But (conclusion from the cases of knives and candidates): It seems that there are expressions (concepts) that have purely *descriptive circumstances* of application, but (intuitively) *evaluative consequences* of application, hence which embody or incorporate inferences from descriptive premises to evaluative conclusions. Are such expressions or concepts (often the ‘thick’ moral concepts that Bernard Williams and ‘non-centralists’ like Susan Hurley champion, concepts such as cruel, and considerate):

- i. Semantically defective, precisely because they embody inferences from ‘is’ to ‘ought’ (the “naturalistic fallacy”)? [Lay out Dummett’s ‘**Boche**’ example (of an inappropriate pejorative used by the French in the first World War, hence *perhaps* sufficiently removed from us that we can think about it: circumstances of application is that one is German, consequences of application that one is barbarous or more prone to cruelty than other Europeans.)] Or
- ii. Absolutely crucial in providing the links between how things are and what we should do that make practical reasoning possible in the first place? (Of course, on an error theory, they might be *both*.) [Say how one can block the argument that ‘Boche’ shows this is wrong: by seeing it as incorporating an inference that is indeed *bad*, but bad because it has counterexamples, not because of its very form.]

g) Notice that Jackson aspires to use something like *descriptions of desires* as *his* amphibious link between description and evaluation—at least when combined with both counterfactual conditional alethic *modal* elements (what one *would* desire if...) and an *ideality* condition, about which I’ll express some skepticism later.

h) But if we accept that there can be concepts with descriptive circumstances of application and evaluative consequences of application (or equivalently—at least

according to us inferentialists about conceptual content—*inferences from descriptive premises to evaluative conclusions*), then what should we mean by ‘descriptive concept’ (expressed by some bit of “descriptive vocabulary”)? Do these count?

i) t

6. Is *describing* or *descriptive vocabulary* semantically autonomous? Is it *pragmatically* autonomous? That is, could there be a language in which *all* one did was *describe*? (Use vocabulary descriptively, use descriptive vocabulary.) If not (and Sellars will argue that *explaining* and *describing* are two sides of one coin—though *much*, but not *all* of the vocabulary used in *explaining* is and must be *descriptive* vocabulary).

7. Putting aside for the moment the question of how we move from the notion of description as the *activity* of *describing* to description as a sort of *content*, expressed by a distinctive *vocabulary*, we can ask: How do we tell whether what we are doing is *describing* (using vocabulary descriptively, using descriptive vocabulary [assuming, for now, that we *can* make the transition from the first, force notion, to the second, content notion])? Thus non-cognitivists like Blackburn and Gibbard distinguish taking a term like ‘good’ as *describing* or as they often say *reporting* (which I don’t like, since I want to reserve the latter term to mark the distinction between *non-inferential* descriptions and *theoretical* ones, which are *not* reports) something objective or outer, on the one hand, and *expressing* something subjective or inner, on the other.

- a) This question matters, because a favorite ploy of philosophers is to take puzzling vocabulary *out* of the descriptive realm. This has been tried for ‘good’ (where the line is that it is not used to *describe* but to *praise*), and for ‘true’ (where the line is that it is not used to *describe* but to *endorse*), ‘beautiful’ (where the line is that it is not used to *describe* but to *appreciate* or *admire*). So the “non-descriptive speech-act” ploy has been used for *all* of the Good, the True, and the Beautiful.
- b) Geach (in his gem-like ‘Ascriptivism’) asks: what are the limits of this ploy? What are the rules of the game? ‘Macarize’ shows that we *could* run this line for *any* expression, even paradigmatically descriptive ones.
- c) Response (the “Frege-Geach test”): the embedding test, for forming a semantic content without attaching a pragmatic force—a semantic content that is *truth-evaluatable* and in *that* sense descriptive. “Disciplined syntacticism.” (I don’t like this term, because what is a declarative sentence in the relevant sense is a *semantic*, not a *syntactic* matter, even though the linguists also have a notion of declarative sentence that *is* syntactic. (Though by and large, the notion of a sentence is taken for granted as an *input* to syntactic theorizing.) So I will use ‘*declarativism*’ instead.)
- d) This countermove will protect description (the act of *describing*, and *maybe* therefore descriptive vocabulary) from *corrosive macarazation*. But in *this* sense of ‘descriptive’—truth evaluable—moral claims are *already* descriptive. Using the form of *declarative sentences* to pick out the descriptive makes the task of the descriptivist in ethics too easy.
- e) So it seems we need something in between: some way of picking out descriptive vocabulary that is *not* merely a matter of declarative sentences, but will also block the danger of *corrosive universal macarazation*—packing *everything* into the pragmatics, by multiplying speech acts for each predicate.

f) Q: What does Jackson have to offer? A: Some paradigm cases, the claim that things “like that” (presumably: in relevant respects, namely...?), and the generous offer that anything that seems at all suspect should be treated as non-descriptive. So only the *clear* cases are to get in.

g)

h) t

8. Suppose there is such a thing as descriptive vocabulary, and a distinctive range of facts it can be used to *state*. (Or is this just *all* facts? That is the minimal *declarativist position*.)

a) Is *modal* vocabulary a kind of *descriptive* vocabulary? If so, then the description of *each* world makes reference to *all* the other (accessible) worlds. Sellars argues, as we'll see, that modal vocabulary is *not* descriptive. It makes explicit *explanatory* commitments and connections, and its job is *not* to describe how the world is. But if modal vocabulary is *not* descriptive, then the descriptive vocabulary is too impoverished even to describe dispositions and nomological connections; it is too impoverished to do physics in. Why should we think the ethical, or anything else interesting, supervenes on *it*?

b) Is ‘true’ a bit of descriptive vocabulary? In saying that some sentence is true, am I describing it, or evaluating it? (Cf. ‘good knife’) For any declarative sentence *p*, including all the descriptive ones, *p* is true in any possible world just in case “It is true that *p*” is true there. This is a necessary truth, and it is *a priori*, because analytic of the meaning of ‘true’. So if we say that ‘true’ is *not* descriptive, we have to say that ‘descriptive’ is highly intensional. For some non-descriptive phrase is necessarily and *a priori* equivalent to each descriptive one.

c) What about *psychological* vocabulary. When I say what someone believes or desires, am I describing her? The answer had better be ‘yes’ (indeed, for modals too), since FJ’s own account in Chapter Six appeals to what we *would desire* if...—indeed, under “ideal” conditions, itself a condition he is committed to being descriptive and not evaluative (at least not *ethically* evaluative). But then what about saying that someone believes that killing is morally wrong? Am I describing her when I say *that*? If not, it is not at all plausible that moral facts supervene on descriptive ones, since moral *attitudes* are *surely* real and important to the existence of such facts.

d) But now it looks as though we have trouble when we put these claims about *psychological* descriptions together with the claims about *semantic* descriptions. For what about saying “S believes that killing is morally wrong, and her belief is true”? Is *that* a description, first of the believer and then of her belief? If so, descriptivism in ethics can indeed be bought cheaply.

e) The only way out I can see is to say that ‘true’ is *not* descriptive if what it applies to is not descriptive. This is a tough line to take: ‘true’ means true. Is ‘true’ *ambiguous*? Does *it* take a different sense depending on what it applies to? But in any case, of course descriptivism claims in the end that moral evaluations *are* descriptive. So it would seem that describing them as true *is* describing them. So this reply only is available on the counterfactual hypothesis that descriptivism is false. (By the way, is *descriptivism* a descriptive hypothesis? It makes essential reference to *moral vocabulary*. Are we *describing* something when we call it “moral vocabulary”?)

## 9. Complete descriptions

- a) The moral facts do not supervene on the descriptive facts as specified in every descriptive vocabulary whatsoever. For I can have a descriptive vocabulary that is so expressively impoverished that all one can do is distinguish between things that are have a mass greater than or equal to 1 gram, and the rest.
- b) So we need a sufficiently expressively powerful descriptive vocabulary. In effect, FJ is Ramsifying, asserting an existential quantification with the vocabulary in question replaced by a variable: There is some descriptive vocabulary *D* such that the facts as expressed in *any* (not just some, because again, we could have a very expressively impoverished moral vocabulary, distinguishing only acts that were identical to my good deed of yesterday, and the rest) moral vocabulary supervene on the facts as specified in *D*.
- c) Now that this vocabulary is a *descriptive* vocabulary is a property it has—presumably, a *descriptive* property. What property is it? (For that is part of the content of the Ramsification, what constrains the realizers that can make the existential claim true.) As we've seen, FJ doesn't really tell us. In effect, he Ramisifies on *this* property, *too*. That is, he tells us that there is something to mean by descriptive such that there is a vocabulary that is descriptive in that sense such that the moral facts as specified in any moral vocabulary supervene on the facts specified in the vocabulary that is descriptive in that sense.
- d) And when in Part II we look at moral functionalism, we will see that he not only sees moral properties as functional properties specified by a Ramsified theory, but that even the non-moral properties that are held fixed in *that* Ramsification are in effect themselves Ramsified: moral functionalism just tells us that *there are* non-moral descriptive predicates (hence properties) such that *there is* a theory formulated using them *and* (some) moral predicates that are implicitly defined (cf. Beth) by the Ramsification of that theory that eliminates all the moral predicates. (FJ's responses to objections, and the variety of issues he claims to be entitled to be noncommittal about, show that he is Ramsifying in this way and to this extent.)
- e) What is left that is *not* Ramsified away in the final supervenience claim, that gives it its content and substance? (Remarkably, given the *extremely* abstract, functional, almost purely structural character of the claim, FJ still manages to give a *bad* argument for it. (The one discussed in (2).)

## 10. Part II: On Jackson's particular proposal for a descriptive account of moral properties, in Chapter 6:

- a) It is to be extracted from the folk theory of morality, when we consider what that theory will or would converge to in achieving ideal reflective equilibrium (the "mature" folk theory).
- b) Jackson is committed to this being a *theory*, since moral functionalism tells us that the Ramsification of this theory is what implicitly defines (in either a meaning-giving or a reference-fixing way, FJ wants to be noncommittal on that topic). So any *practical wisdom* (phronesis) or *ability* to apply moral concepts in concrete circumstances is *not* permitted to be part of the meaning of moral concepts, except insofar as it can be *explicitly formulated* as a *theory*. Practice must take the form of principles, knowing-how rendered *exhaustively* as knowing-that. To allow any such practical abilities (the result of being properly brought up or acculturated—gebildet)

to serve as essential elements of the articulation of the content of *any* moral concepts would make the Ramsification proposal unworkable. So this is not just moral *functionalism* about the contents of moral concepts, but *theoretical* functionalism about those contents. Jackson just makes the assumption that this commitment is OK; he says *nothing at all* by way of justifying or entitling himself to it.

- c) The “mature folk theory” of moral concepts, to which it is presumed *actual* folk theories are *would* converge (if some *ideal* set of conditions—which must be specifiable in *non-moral* terms—obtained) also, it seems, cannot appeal to an in-principle open-ended set of *new particular cases* or situations in which moral concepts must be applied, nor to the effects of new *discussions* of them (moral theory should be different *after* Henry James’s novels).
- d) Is the notion of convergence possibly misplaced here. *If* the issue is a *descriptive* one, and we already have our concepts fully in place, we might think that an appropriate (“disciplined”) epistemology or methodology would guarantee (if things don’t go *too* far wrong) that we would converge on full and correct descriptions of how things are, in terms of those concepts. But here it seems FJ wants to have convergence of the moral *concepts*, in some sense independently of our discovery of the moral *facts* that we would express with those concepts.
- e) What are the concepts that one is allowed to use in the Ramsified theory that implicitly defines the moral concepts? They must be *non-moral* concepts, since we are Ramsifying all the moral ones, in order to specify their functional roles.
- f) Whatever they are, is it plausible that we can specify the contents of moral concepts entirely in non-moral terms, that is, without giving *any* of the moral concepts their *intended* interpretations? What reason do we have to think that *any* and *every* set of predicates/properties that can realize the Ramsey predicate (make true the Ramsey sentence) by being appropriately related to *non-moral* properties *is* a set of *moral* properties?

#### 11. On “moral functionalism”:

- a) Is this view supposed to be peculiar to the *moral* concepts? Shouldn’t we be *physical* functionalists, too? That is, shouldn’t we for *other* vocabularies (aesthetic, etiquette, sociological...) see them as having their content conferred by their role in a theory, which we can Ramsify to specify those roles? (Q: What would remain *unRamsified* in the case of *physical* concepts? *Observational vocabulary*?) Note that this is the ‘theory theory’ (not to be confused with the metatheory in the philosophy of mind that goes by the same name) that Wilson complains about: theoretical terms are implicitly defined by the theories they are part of. Such an account ignores the vicissitudes of the development of content that goes with their actual course of application in concrete circumstances—which is what not only unfolds, but also develops and helps confer their content.
- b) The semantic autonomy issue again: Why should we think that we can do something like this—Ramsify a theory to get an account of the semantic roles that are the meanings (or fix the references)—vocabulary by vocabulary, in the absence of other vocabularies? Can we do it all at once for *all* vocabularies? But then we get just a formal logical structure, with no non-logical vocabulary left *unRamsified*. (Not even ‘cause’, or ‘time’.) And that, absent language entries and exits, is *not* enough to confer non-logical empirical content.

12. FJ's particular proposal, again:

a) FJ suggests that what we ought morally to do is what we would all things considered desire to do under ideal circumstances. (So *ideality* comes into his story about the descriptive character of moral norms *twice*: once in getting the “*mature* folk conception” to Ramsify, and *again* when we look at what one *would* desire under *ideal* circumstances.)

b) Passages:

[159] “On the moral functionalist story, to believe that something is right is to believe in part that it is what we would in ideal circumstances desire, where we can regard the rubric ‘would in ideal circumstances desire’ as covering the possible spellings-out already mentioned—perhaps what we would desire in ideal circumstances is what we would desire when our first-order desires square with our second-order desires, or when our desires square with what we would converge on stably desiring after reflection, perhaps taking into account the desires of our community, or something along these gestured-at lines.”

[157]: “Perhaps the properties it is rational to desire are the ones we would desire to desire after critical reflection on and full acquaintance with them; perhaps they are the properties our ideal selves, possibly in the guise of an ideal observer, would desire, or would desire to desire; perhaps they are the properties our idealized selves would converge on desiring after discussion; perhaps they are the more stable of our long-term desires; perhaps they are the desires we are prepared to universalize in the sense of being the desires we would allow anyone and everyone to act on;...The details will not matter for what follows.”

[157]: “To believe that A is right is *inter alia* to believe that A has the property it is rational to desire. And this surely is a belief that *points towards* doing A.”

[158]: “When you believe that A is right, you believe in part that A has properties it is rational to desire *all things considered*.” [Note that this builds in both *independence* of contingent desires and statuses, and *overridingness*.]

[159]: “What shows that the belief that A is right is in part a belief about what would be ideally desired, is that we form it when it is true that **we would in ideal circumstances desire A.**”

[160] (final page of the final lecture): “**We form the belief that A is right when we are disposed to desire it in ideal circumstances.**”

13. Q: Why is FJ's view a view specifically about *moral* normativity?

- a) Can't I paraphrase *any* ‘ought’ statement into a statement about what one *would* (or, as it seems from the considerations above one might equally well say, given the *ideality* of the ‘would’: *should*) desire under some ideal circumstances?
- b) Thus: “This is a good knife,” → “This is the knife one would ideally (=should?) desire to have.” “Bank employees ought to wear neckties,” → “Under ideal conditions, bank employees would desire to wear neckties,” or “Under ideal circumstances, we would desire that bank employees wear neckties,” and so on.
- c) Such examples may seem to offer something corresponding to *hypothetical imperatives*: they are conditioned in some way by needing a knife, or being a bank employee.

- d) Q: But how is FJ thinking about what is distinctive of *moral* ‘ought’s? ☶A: Is it that they correspond to *categorical* imperatives, that is, are *not* conditional on one’s other contingent desires (to cut) or status (as a bank employee)?
- e) But there are at least two importantly different ways to hear that:
  - i) It could be that this means that independently of any contingent desires or statuses [and how would we go about filling in this list?], one has reason to do (would ideally or should) desire to do) what thereby counts as what one morally ought to do.
  - ii) Or it could mean that independently of all that contingent stuff, *all things considered* this is what one would have the *most* reason to do (would ideally or should) desire *overridingly* to do.

The difference is that on option (i) we leave open the possibility of *weak* moral reasons/desires/‘oughts’, ones that might properly be overridden by *strong* prudential or other kinds of reasons/desires.

- f) One of the things *natural-kind skepticism* about the moral is skeptical about is whether we have any (non-theological, non-apologetic, non-utopian) reason to think there *is* any range of reasons/desires/‘ought’s having the property described in (e-ii).
- g) What if we are natural-kind skeptics about the moral? Then we take moral reasons at least often to be genuine reasons for action, but do not take it that there is anything that a distinguishable set of reasons has in common in virtue of which they deserve to count as *moral* reasons.

14. There are at least two fundamental issues here, which interact:

- i) Is there any way of specifying the *ideality* conditions that is both *non-circular*, and *plausibly gives the right result*, i.e. is *materially adequate*?
- ii) What is the *motivational force* of judging that I would under ideal conditions desire that *p*? That is, describing something as what I would (but do not now) desire under counterfactual conditions must serve as a motivating reason to bring it about that *p*.

- a) Now the fact that under *some* counterfactual circumstances I would desire that *p* does *not* serve as a motivating reason for bringing it about that *p*. For that is plausibly true for *any p*.
- b) So it must be something about the *ideality* of the conditions that makes the fact that under *those* counterfactual conditions I *would* desire that *p* give me a reason, capable of competing with what I *actually do* desire, to bring it about that *p*.
- c) So the challenge is to give a non-circular, plausibly materially adequate account of the sense of *ideal conditions* that will provide *motivating reasons* for wanting *now* to do what I *would* want to do under those conditions.
- d) Given the right notion of ideality, in some sense, I *should* do what I *would* want to do under those ideal counterfactual conditions. What is the status of that ‘should’? It better not be *moral*, or the account will be circular. Is it prudential/instrumental? But prudence is a matter of acting so as to get what one wants or prefers. But that is what one *actually* wants or prefers. We could define a notion of *ideal* prudence, which is getting what one *ideally* wants or prefers. But if we appeal to *that* notion in defining the ‘should’ above, we go around in a very small circle, and so violate non-circularity.

- e) So it is a criterion of adequacy of the definition of the ‘ideal’ conditions (the ones in which I *would desire* to do what I *morally ought* to do) that it can be specified *without* the use of specifically *moral* vocabulary.
- f) What should we say about the auxiliary hypotheses that *can* be appealed to in defining that notion of *ideality*? We can, I think, appeal to *cognitive* ideality (though not, of course, ideal conditions for knowing what one ought morally to do)—for instance by requiring that I have only *true* collateral beliefs (and perhaps, not be *ignorant* of crucial features of things), so long as we don’t allow that the true beliefs (or the things I must not be ignorant about) concern *moral* states of affairs.
- g) Suppose I have a neurosis, or an overriding commitment to The Cause, that goes so deep into my personality that pretty much no matter what counterfactual circumstances *I* (with my at least somewhat essential neurosis) was put in, *I* would not, in fact, desire to do what intuitively I morally *ought* to do? Would that perversity of my motivational structure mean that it is not in fact the case that *I ought* morally to do it? Or, if we need to block that outcome (because of implicit constraints on getting the folk-theory of moral facts right), is that to be done by abstracting away from those peculiarities of mine, making their elimination part of the ‘ideal counterfactual circumstances’ we are considering? In that case, what do they have to do with *me* and *my* motivating desires? Or is it enough that *most* people, or *more rational* people *would* have the desire in question, though I am outlier who would not? In that case, again, what is that fact supposed to have to do with giving *me* a motivating reason, understood in terms of *my* desires?

15.

- a) Notice that FJ assumes that motivation is by *desires* (even though this doesn’t seem to help when it is only *possible* or *counterfactual* desires that are available). The idea of being motivated by one’s acknowledgment of a norm, value, or commitment, apart from one’s desire to conform to, pursue, or fulfill it, is not part of his setting.
- b) That is, FJ seems to assume that *descriptivism* about norms (recall the question of whether it really is specifically *moral* norms that really are at issue for him, in spite of what he says) demands *humean motivational internalism* about norms: that I can only be motivated by norms that I *desire* to conform to or satisfy. He does not consider *kantian motivational externalism* about norms: that I might be motivated by what I have *reason* to do, by norms I acknowledge as binding, quite apart from my desires. [If there is time, do a quick riff here on how the fact that if I have strong desires that *conflict* with the norms I acknowledge I might not in the end be successfully motivated by the norm or reason does *not* show that in the absence of strong countervailing desires I can act on a reason or acknowledged norm *only* if I *desire* to do so.]
- c) In fact there is a contradiction (at least a pragmatic/methodological one) here. For, on the one hand, FJ is assuming a model that takes it that *only* desires can motivate. On the other hand, he is claiming that the mere contemplation of, or judgment that, under certain counterfactual circumstances one *would* desire something can motivate one.
- d) But if the latter is true, then why couldn’t one directly be motivated simply by acknowledging the applicability of a norm or a commitment?

e)

Afterword:

The current state of play w/res to the Master Argument of Jackson's Chapter 5:

16. The argument is for the conditional: if ethical vocabularies/facts globally supervene on descriptive ones, then ethical properties are descriptive properties—in a sense rich enough to support strong versions of both sides of classical Carnap-Nagel reduction (definability of concepts and derivability of ethical from descriptive facts). Here is the argument (almost) *in toto*:

[122-3] E is a sentence framed in descriptive and ethical vocabulary.

“Each world at which E is true will have some *descriptive nature*...And for each such world, there will be a sentence containing only descriptive terms that gives that nature in full. Now let  $w_1, w_2$ , etc. be the worlds where E is true, and let  $D_1, D_2$ , etc. be purely descriptive sentences true at  $w_1, w_2$ , etc., respectively, which give the full descriptive nature of  $w_1, w_2$ , etc.. Then the disjunction of  $D_1, D_2$ , etc., will also be a purely descriptive sentence, call it D. But then E entails and is entailed by D. For every world where E is true is a world where one or the other of the  $D_i$  is true, so E entails D. Moreover, every world where one or the other of the  $D_i$  are true is a world where E is true, as otherwise we would have a violation of (S) [the supervenience of the ethical on the descriptive]. Therefore D entails E. The same line of thought can be applied *mutatis mutandis* to ethical and descriptive predicates and open sentences: for any ethical predicate there is a purely descriptive one that is necessarily co-extensive with it. **It follows that ethical properties are descriptive properties.”**

17. I claimed that the most interesting philosophical issues in the vicinity concern the notion of descriptive term that is presupposed here. For it raises issues different from, and deeper than, those (weighty and worrisome enough) raised by supervenience claims that appeal to “the language of physics” or (worse) “the language of natural science.” To recall a few of the central ones:

a) It is not easy to say what *describing* is, or even what *kind* of speech act it should be thought of as. What other speech act-kinds belong at the same level, or contrast with it in the same family? The most popular candidate (and the one most relevant to Jackson's argument) is *evaluation*. But these don't seem to be mutually exclusive in any straightforward sense, as the plethora of terms with descriptive circumstances of application and evaluative consequences of application (and the corresponding inferences) shows.

b) Downstream of that difficulty is another: Even if we settle what the activity of *describing* is, we need an account of and rationale for the definition of a notion of *descriptive term* (descriptive vocabulary) in terms of it. Is this a way of classifying types, or tokens? Is a descriptive term, for instance, a linguistic expression that *can* be used to describe, or an expression that can *only* be used to describe? Are we sure that the first option excludes anything, or that the second includes anything? Maybe all or most expressions *can* be used to describe and all or most expressions that can be used to describe can also be used to do other kinds of things. If this is the wrong way to go about moving from a kind of speech act to a kind of expression or concept, what

is the right way? (Talk of the “vocabulary of physics” or of semantics, etiquette, and so on does not raise *this* sort of problem.)

c) Geach’s ‘macarize’ shows that for any intuitively descriptive term (e.g. ‘happy’) we *can* introduce a characteristic speech act, distinct from description, that one performs just by applying that term to something. The notion of description will be emptied of significance if not defined broadly enough to block the promiscuous use of this ploy: to block corrosive, rampant macarization. Geach’s own suggestion is what I called ‘declarativism’ (and Jackson, on Wright’s behalf, “disciplined syntacticism.”) In the present context, such a view identifies being descriptive with being usable to form declarative sentences, hence to be potentially fact-stating, and true-evaluative, as determined by the capacity of such sentences to appear as content-specifying components in compounds, paradigmatically as the antecedents of conditionals or in ‘that’-clauses. The strategy for blocking rampant macarization is admirably clear, and does the job. But it leaves no room for a question about the relation between describing and evaluating, and hence about the relation between descriptive and evaluative vocabulary. By its lights, saying that an act is good or right, or that an agent ought to do something *is* stating facts about them, hence describing them, attributing properties to them. If there is to be an intelligible issue of the sort Jackson’s thesis addresses, *descriptive* terms have to be picked out some other way.

d) One reason we need to find something good to mean by ‘descriptive’ is in order to address the issue Sellars will raise (in “Counterfactuals, Dispositions, and the Causal Modalities): whether there can be such a thing as a purely descriptive language. We might paraphrase this as the question of whether describing is potentially an *autonomous* discursive practice: a language-game one could play though one played no other. Or is it necessarily a *fragment*, intelligible only as an aspect of a wider discursive practice, encompassing other kinds of speech acts, and (so?) other kinds of vocabulary? Is descriptive vocabulary *semantically* autonomous, in the sense that the meanings of expressions of this kind can be made sense of apart from their relations to *other* kinds of vocabulary? (The issues of pragmatic autonomy and semantic autonomy are intimately related—particularly if descriptive vocabulary is to be picked out in terms of the speech act of describing. But they need not in general be taken to be the *same* issues.) *Descriptive*, as opposed to *scientific*, naturalism, paradigmatically about at least some kind of *normative* discourse, is very difficult to formulate so as not to involve commitment to some such autonomy thesis. (Sellars denies the semantic and pragmatic autonomy of descriptive vocabulary, and hence rejects “descriptivism”, while connecting scientific and descriptive naturalism through the *scientia mensura* principle: “In the dimension of describing and explaining, science is the measure of all things, of those that are, that they are, and of those that are not, that they are not.)

18. I had hoped to be able to put Jackson’s Master Argument to one side without having to enter into these dark thickets—even though exploring those thickets is, as I see it, the most philosophically interesting enterprise in the vicinity of descriptive naturalism, and the one we are plunging into next. Jackson is prepared to concede that both the conjunctive  $D_i$  specifying the “full descriptive nature” of each world  $w_i$ , and the disjunction of them all  $D$  will have to be formulated in some infinitary (though “purely

descriptive") language, and further, that neither the conjuncts nor the disjuncts of them need be even recursively enumerable. I pointed out that things are much worse than that. He is being disingenuous in indexing the disjuncts. How many possible worlds are there? I don't know. (I don't even know exactly what the question means. It probably depends on what use we are planning to make of the notion.) But only on very special assumptions—and at the cost of considerable difficulty later on in employing the notion of possible world in doing intensional semantics, articulating counterfactuals, and so on—can there be taken to be only a *countable* infinity of possible worlds. (Whether a countable set of sentences can be thought to express the "full descriptive nature" of a possible world is a question that cannot sensibly be addressed without worrying about the sort of issue raised in (2) above, as we shall see.) And that means that the 'language' in which the disjunction-of-conjunctions D is formulate will have to be a more than countably infinitary language. So the 'open sentences' expressing the descriptive properties equivalent to ethical ones may be formed by omitting uncountably many (descriptive) terms, to be replaced by variables that can later be bound by quantifiers. There are not enough sub-atomic particles in this (possibly in any possible) universe to write down or otherwise be such 'expressions'. [OK: The warning-light on your MADE-UP PHILOSOPHERS' PHYSICS alert system should be going off.]

19. I took those considerations to be enough reason not to take Jackson's argument seriously. But it was pointed out correctly that there are perfectly formally well-behaved infinitary languages that permit conjunctions, disjunctions, and quantification applied to collections of sentences of arbitrarily high cardinalities. Jackson doesn't acknowledge that his argument goes through only in such a context, or that the descriptive properties he constructs are of this character, but so what? He is making a metaphysical point. He has already conceded that the descriptive equivalents he shows must exist for ethical claims and predicates are completely intractable in any practical terms—being infinitary and messily gerrymandered at least in the sense of not being recursively enumerable. So what difference does this further concession make?

20. I think this response does show that I cannot dismiss Jackson's argument at a stroke by invoking the cardinality of the set of possible worlds that he blithely indexes his disjunctions-of-conjunctions to. And that means that in assessing it we cannot avoid delving into the messy considerations about what descriptive terms and vocabularies are. What we see when we do that, I think, is a mismatch of frameworks. The seriously infinitary languages we need to invoke to formulate Jackson's descriptive equivalents of normative facts and properties are complex set-theoretic constructions. (That is why the fact that there don't begin to be enough *physical* things to be *expressions* in that language doesn't matter for *that* sense of 'language'.) It is (to underestimate the point) not easy to see how to bring languages in this formal sense into contact with concerns about *acts* of describing, or the pragmatic or semantic autonomy of "purely descriptive languages." It seems like a category mistake to ask how one would need to *use* expressions to make them mean what the sentences of these "purely descriptive" abstract set-theoretic 'languages' mean. But going into this question requires resources well beyond what I was bringing to bear in my original argument.