

### **Gordon on Categorical Imperatives, Possible Worlds, and Self-Presupposing Desires**

As I said in my e-mail, I now think there's a problem about Gordon's formulation even of the future-health example.

He says, "My present desire to be healthy at eighty is very clearly a desire that I be healthy at that age *whether or not* I should happen *then* to desire to be healthy" (p. 106). Though that sounds right, it puts the point in terms of the desire's semantic content: "My desire that... *is*... a desire that..." But the two complement clauses are not obviously logically equivalent. If we individuate desires according to their contents, we are presumably talking two distinct desires, not a desire with one content that *is* also a desire with a distinct content. Assuming closure failure, of course even logical equivalence of contents doesn't entail same desire. In fact, even closure wouldn't necessarily mean same desire.

(But are the complement clauses after all logically equivalent? The semantic content of the second desire clearly entails that of the first. But does the content of the first entail that of the second? In favor: "Duh. If I will, in fact, be healthy at eighty, then trivially I will be so whether or not *anything* except what's entailed by my being healthy at eighty; the 'whether or not' 'stipulation' or assurance is triflingly redundant." Against: "A desire's content is always in a way conditional [cf. our discussion last week of the fine-grainedness issue: "I didn't mean *that* way (famous for that kind of thing)"]. When we ascribe the desire "that P," "P" abbreviates something a bit more complicated: that P assuming certain defaults hold. The content is not really: that P, period, no matter what."

I think the latter view is problematic. It seems to me to trade on understandable mismatches between verbal desire ascriptions and either what the speaker really means or what the subject really desires. Let's continue to think about this. But for now, the point is to seek a less dubious description of the future-health example.)

One description is counterfactual: "*Would* you still (now) desire to be healthy at eighty if you now knew that when you do turn eighty you won't then desire to be healthy?" That's a perfectly good question, answered yes. But it doesn't extrapolate to the present-health case ("*Would* you still desire to be healthy now if you now knew that you don't desire to be healthy?"—trivially, no).

Or, possibly the desire really is conditional in content: It's really *to be healthy if [default conditions] and even if I will then not any longer desire to be healthy*. That extrapolates, at least without contradiction, to the present-health case; I can desire *to be healthy now if [default conditions] and even if I do not now desire to be healthy*. Indeed, that sounds like what Gordon means. But this idea of interpreting-as-conditional would apply to most any desire, since "even if...?" questions can always be raised. Is (practically) every desire merely conditional? I don't think so. I want another cup of coffee now. "Don't you mean, you want *a cup of coffee if your drinking one won't cause a nuclear holocaust?*" No (though I do also want that), because I believe that drinking

one won't cause a nuclear holocaust and I quite unconditionally want the damn cup of coffee, even if I *would not* want that if I believed differently.

Maybe we should reconsider the policy of individuating desires by distinctness of content, because we've already considered the idea that a single desire may have more than one content. If Stampe is right in assimilating desires to perceptual states and I and Peacocke and Noë are right in maintaining that a single perceptual state has multiple layers of content, that may be the right way to go. But it would be a new kind of layering, for a single desire to have both a straightforward unconditional content and a complex conditional one.

### Possible worlds

As we saw, Gordon can't mean *relative to* by "relative to" (p. 107). I think what he means is that with every desire there is *associated* a set of worlds. It's clear that the set he means is not the standard Hintikka content set. In some of its worlds the desire is satisfied, but in others it's frustrated. The set excludes worlds in which the desire is neither satisfied nor frustrated. (Gordon doesn't say that the set contains *all* the satisfaction/frustration worlds, but I'm taking it that that's what he means.)

There's a crucial ambiguity: On the one hand, Gordon says that a desire is satisfied at a world when ("i.e.") its satisfaction condition is true there, and frustrated at a world when its satisfaction condition is false there. On the other hand, he says that at a world not containing the desirer, the satisfaction condition lacks truth-value. But what, then, of my desire for world peace? Take a world at which I don't exist but in which there is world peace. At that world, the *semantic* satisfaction condition is true, but Gordon says the satisfaction condition is truth-valueless. So presumably he means something stronger by "satisfaction condition" than the mere semantic one.

The next obvious thing for him to mean by it would be the condition under which *the desire itself gets satisfied*, i.e., the desire exists and also is satisfied rather than merely being extinguished. (The Lycanless peaceful world is not one of those, because my desire doesn't exist there.) But Gordon can't mean that either, because if he did he couldn't raise his "main" question of whether worlds in which one does not have the relevant desire qualify to be in the associated set.

Of course, his position is that for the categorical-imperative-type desires, some no-desire worlds do so qualify. That's his possible-worlds way of saying what he was trying to say about the present-health example. For the merely appetitive desires, no-desire worlds are excluded from the associated set. But those claims don't help unless we understand what set the set is supposed to be in the first place.

I judge that pursuing this matter further is not a good use of my time. The possible-worlds formulation is, at best, underexpounded and not clear enough to prove "helpful" (p. 107). Also, for the reasons given above, I suspect that what Gordon is trying to say is not coherent. It would be better just to turn to the examples on p. 108.

### The examples on p. 108

1. If I cause my eating desire to be extinguished without eating, the desire is not satisfied. 2. Nor is it frustrated, even though I do not eat. 3. Even if I do that and then

(for whatever reason) do eat, the now extinct desire is not satisfied. 4. If I eat but it fails to quench the desire, the desire is not satisfied. These claims seem right, at first.

Gordon takes them to show that the “object” (p. 109) of the desire is not simply *to eat*. He suggests that the object is that one eat *and by doing so* quench the appetite. But now let’s distinguish between hunger, a drive and/or feeling, and a desire qua propositional attitude. (As always, one may desire to eat even though one isn’t hungry, and vice versa.) 1-4 are plainly true if “hunger” is substituted for “desire.” However, what happens if we focus on the notion of “desire to eat” strictly so called, as opposed to hunger? 1 remains true. 2 may be true in a sense of “frustrated” yet to be specified, but the sense will have to be stronger than that of semantic non-satisfaction. 3 may be true but only in a sense stronger than that of semantic satisfaction. 4 depends: In one sense the desire would be satisfied, because what I wanted was to eat and I did eat; it’s just that now I want to eat more.

De Sousa’s time issue rears its head here. An “I-desire,” he said, is “for something *right away*” (p. 87), which suggests that 3 is true after all but leaves 4 up in the air.

There’s also the issue of fine-graining of content. You say you desire to eat, so I force-feed you twenty Brussels sprouts, each covered in mold. “I didn’t mean *that way*.”

I don’t know what we should think at this point about desire content. But let’s turn to Gordon’s issue about desires that presuppose their own existence, and look at it more directly.

### Do some desires presuppose their own existence?

Gordon says that that formulation is short for the desire’s “satisfaction condition”’s presupposing that one has the desire. If he’s right that the desire’s “object” (= the “satisfaction condition”??) is that one eat and by doing so quench the desire itself, that’s a sense in which the desire presupposes its own existence. Contrast the health case: My present desire for health doesn’t have any sort of self-quenching character.

Beth’s reasoning (pp. 110-11) is indeed bad. If her desire to eat presupposes its own existence, that would help explain why the reasoning is bad: Part of the “object” of her desire is to quench the desire itself, and she has no additional reason (such as pleasure) to eat, so she has no reason to be thankful for the desire, and it would not be bad for her to take the quenching pill even though the pill would not *satisfy* the desire. (But I don’t understand what Gordon says about his (1’) and (2’). What seems to be true is only that if (1’)’s supposition were true, Beth would have no reason to believe (2’).)

Actually I’m not sure Gordon means that Beth would have no additional reason to eat. What he says is that she “sees no reason to eat now, other than to satisfy her hunger” (p. 110), and we still don’t know what he means to build into “satisfy” as applied to hunger.

The “fishing” argument at the bottom of p. 111 is very clearly fallacious. So is its desire counterpart: “It’s fortunate that I desire to fish. For if I didn’t, I probably wouldn’t go fishing very often—which would be too bad, because I so much desire to.” That does contrast with the health case, since it’s not only because I desire health that a decline in my health would be bad. But as we discussed in class, what seems to be going on here is not something special about the desire itself, but the intervention of another modality, my

health having *value* of a desire-independent kind. As Anabella noted, Gordon may think he can reduce such value to categorical-imperative-type desires (pp. 106-07), which would be an interesting project, but he doesn't undertake that here.

### The explanation

How does all that, if correct, explain the original distinction between desires that close the circle and those that don't? I gather Gordon's answer is that Alf's reasoning is good and shows why Alf's desire for health is itself a practical reason for having it, while Beth's reasoning is bad and so can't show that her appetitive desire is a reason for having it. That's fine so far. But notice that Alf's argument requires premise 2, about *badness*. It's again because of that independent value that his desire for health serves as a practical reason for itself. I don't think that's an objection to the explanation, but I'm not sure.

As you can tell, I'm not confident that I understand what Gordon is up to. If you can find a clearer interpretation and defense, that would be an excellent project.