WALTER SINNOTT-ARMSTRONG

SOME PROBLEMS FOR GIBBARD'S
NORM-EXPRESSIVISM

In *Wise Choices, Apt Feelings*,[1] Allan Gibbard discusses a wide variety
of topics, from rationality and morality to emotions, representation,
objectivity, and epistemology. He also employs a wide variety of
methods drawn from biology, psychology, pragmatism, and decision
theory. Gibbard wields all of these tools with great skill and says
insightful things about every topic he mentions. Nonetheless, this is a
session where an "author meets critics", not fans, so I will spend most
of my time as a critic.

The central claim in Gibbard's book, the claim which motivates and
ties together the other strands, is Gibbard's norm-expressivist analysis.
This analysis openly falls into the tradition of emotivism. Emotivism fell
out of favor when it ran into some serious problems. Gibbard is fully
aware of these problems and carefully formulates his analysis so as to
avoid or solve each of them. Nonetheless, I will argue that Gibbard's
revisions of traditional expressivism fail, because they either create new
problems of their own or surrender what is distinctively expressivist
about his analysis.

1. EXPRESSIVISM

The driving force behind Gibbard's theory is his naturalism. (8, 35) The
main goal of his analysis is to explain normative thought and talk as a
part of nature. (32–3) His subject is "ordinary" or "everyday" (32,
253) normative thought and talk.[2]

This project lies squarely within the fields of philosophical semantics
and psychology. Gibbard also happens to deny that there are any
normative facts, but that ontological claim is separate from his semantic
analysis. One could reject his semantic analysis and still accept his
ontological claim that there are no normative facts. What will concern

me here is not his view about what there is but his views about what we mean and believe.

When Gibbard tells us what normative judgments mean, he does not lay out their truth conditions. Instead, he specifies the speech act which they perform:

> The analysis is not directly of what it is for something to *be* rational, but of what it is for someone to *judge* that something is rational. We explain the term by saying what state of mind it expresses. In this sense the analysis is *expressivistic*. . . . The analysis is non-cognitivistic in the narrow sense that, according to it, to call something rational is not to state a matter of fact, either truly or falsely. (8)

This passage contains two separate claims. Gibbard's positive claim is that normative judgments *do* express special states of mind (which he will specify later). His negative claim is that normative judgments do *not* state (normative) facts or attribute (normative) properties (9, 105).[3]

It is important to realize that Gibbard's negative claim does not follow from his positive claim. One could agree that to assert a normative judgment is to express a special state of mind and also think that normative judgments simultaneously state normative facts. If so, Gibbard's positive claim is true, but his negative claim is false.

Since these two claims are independent, there are four possible positions. Pure expressivists make Gibbard's negative claim as well as his positive claim. Pure descriptivists deny both claims. Mixed theories see normative judgments as a mixture of expression and description, so they accept Gibbard's positive claim but reject his negative claim. Negators think normative judgments are neither expressive nor descriptive, so they deny Gibbard's positive claim but accept his negative claim.[4]

Given these alternatives, Gibbard needs to argue not only against pure descriptivism but also against the other alternatives, especially mixed theories.[5] Unfortunately, Gibbard's arguments work only against pure descriptivism and not against mixed theories. For example, when he first introduces descriptivism, he defines it as the claim that "if a person calls something rational, it would be best to hear him as describing it, as ascribing a property to it." He quickly rejects this claim on the ground that "It misses the chief point of calling something 'rational': the endorsement the term connotes." (10; cf. also 33–4) The first problem with this argument is that even pure descriptivists can

agree that the term 'rational' is *usually* used to endorse, and Gibbard never shows that the term 'rational' in this sense must *always* be used to endorse (or to express). Furthermore, even if all normative judgments must be endorsements, the claim that judgments of rationality do ascribe properties does not imply that they do not also endorse actions. Gibbard has simply overlooked the possibility of mixed theories.

This oversight would not be very damaging if mixed theories were mixed up. But they're not. In fact, mixed theories have a lot going for them. Since they agree that normative judgments express special states of mind, they can also explain what Gibbard explains with his claims about expression. When Gibbard says that people who disagree about what is rational are expressing incompatible states of mind, mixed theorists can offer the same explanation.

But there is more than this to explain. Any analysis of normative judgments must also explain their role in arguments, such as modus ponens. This explanatory need is part of what drives Gibbard's opponents to add descriptivist elements to their analyses.

Gibbard recognizes this problem. He first presents his analysis as the claim that to call something "rational" is to express one's acceptance of norms that permit it (7) and to think something is rational is simply to accept norms that permit it (46, 55, 81) or a system of norms that permit it (83, 84). But he admits that this simple, original analysis fails to solve problems of ignorance, communication, naïveté, and, most important, embedding, which includes embedding in arguments. (89–94) These problems require what he admits is "a substantial transformation of the analysis." (92)

2. REVISED EXPRESSIVISM

Gibbard's revised analysis is complex and sophisticated. He still claims that to hold a normative belief is to be in a special state of mind, and to assert a normative statement is to express that state of mind. Now he says that the special state of mind consists in ruling out all combinations[6] of normative systems[7] and total factual possibilities which do not entail the normative judgment. (95–6)

In order to see how this works, we can go through an example. What Gibbard calls the problem of embedding is traditionally raised by

arguments employing modus ponens with normative premises,[8] so let's look at one:

(1)   It is wrong to kill.

(2)   If it is wrong to kill, it is wrong to pay someone to kill.

(3)   Therefore, it is wrong to pay someone to kill.

This argument is obviously valid (even if a premise is false).

Traditional expressivist analyses have trouble with such arguments. Why? The first premise might be used to express some attitude towards killing, but the second premise does not seem to express any attitude towards killing. One can accept the second premise even if one thinks that killing is not wrong. Thus, if the meaning of "wrong" is its use to express an attitude, then the two premises use the term "wrong" with different meanings, so the argument commits the fallacy of equivocation. That makes it hard for expressivists to explain why the argument is valid.

In response, Gibbard offers a straightforward explanation of why this argument is valid. The first premise rules out the set of all combinations of norms and facts in which killing is not wrong. The second premise then rules out the intersection of the set of combinations in which killing is wrong with the set of combinations in which it is not wrong to pay someone else to kill for you. Thus, the premises together rule out the whole set of combinations of norms and facts in which it is not wrong to pay someone else to kill. And this includes every combination that the conclusion rules out. This is supposed to explain why modus ponens is valid.

I think that this works as far as it goes, and it constitutes an improvement on many previous versions of expressivism. However, it still leaves crucial questions unanswered.

First, it is not clear why his revised analysis is "expressivist" in any important way. Compare a parallel analysis in another area: causation. Suppose someone claims that to believe (or to assert) that a causes b is just to be in (or to express) a state of mind which consists in ruling out various combinations of causal laws with non-causal states.[9] The validity of modus ponens with causal judgments would then be explained just as Gibbard explained the validity of modus ponens with normative judg-

ments. However, this analysis and explanation are compatible with a fully realistic view of causal claims as claims about an independent world. Similarly, Gibbard's analysis of the content of normative judgments and his explanation of the validity of modus ponens seem just as compatible with a fully realistic view of normative claims as claims about an independent world. Of course, this analogy does not show that Gibbard's theory is inadequate, but it does show that his theory is not distinctively expressivist so far.[10]

So what *does* make his new analysis expressivist? Two things. First, he still denies that normative judgments have truth values. Second, he claims that to accept a norm is to be motivated in a certain way. I will argue, however, that these very elements which make his theory expressivist also create serious problems.

3. THE VALUE OF CONSISTENCY

A first problem arises because he cannot invoke truth to explain validity. It is not enough simply to define validity so that modus ponens comes out valid. Valid arguments have force because there is something wrong with asserting the premises and denying the conclusion of a valid argument. Gibbard's definitions tell us that these claims are inconsistent (98), but we still need to ask: what's wrong with inconsistency? Descriptivists (including mixed theorists) can answer that inconsistent normative judgments cannot all be true, and truth is the goal of moral inquiry. Gibbard can't say this, as he recognizes, (287) since he denies that normative judgments are true or false. (8) Instead, Gibbard has to introduce a "pragmatic" (290) account of the value of consistency. He claims that when one believes or asserts normative judgments that are inconsistent, such as when one asserts the premises but denies the conclusion of the modus ponens argument above, "the problem is how to do so without opting out of normative discussion altogether, or discovering that I can no longer get others to take my claims seriously." (290) And it is bad to opt out of normative discussion because "we need the benefits of normative discussion" (290; see also 75).

The most important problem with this proposal is that inconsistency does not always bring an end to normative discussion. The audience might not notice the inconsistency. Even if they do notice it, they might

have strong reasons to continue the normative discussion. I don't stop talking to my boss or a friend just because their views are inconsistent. Sometimes inconsistent moral views even seem necessary for success in politics. The point is that the pragmatic costs either do not arise or are overriden in many cases of inconsistency.[11] But there is still something wrong with the inconsistent normative beliefs in these cases. Thus, Gibbard fails to explain what is wrong in all of the cases where there is something wrong with inconsistency. Until he explains this, he has not fully explained the validity of arguments containing normative judgments.

4. NORMATIVE GOVERNANCE

Another problem for Gibbard concerns the mental state that constitutes a normative belief. Gibbard's formalism gives us the content of a normative belief in terms of fact-norm combinations. But that is not enough. As he recognizes (99 ff.), he also needs to specify the relation between a content and a person which constitutes that person believing that content.

One way to do this is to claim that, when a person accepts or believes a normative content, that person is motivated or has a tendency to act accordingly. This kind of internalism is common not only among expressivists but also among their opponents.

The popularity of internalism about motives is strange.[12] I know of no strong argument for it, and counterexamples abound. Judgments of permission are normative, but I can believe that I am (morally and rationally) permitted to go skydiving without being motivated to go skydiving. Judgments about reasons are also normative, but I can believe that I have some (moral or prudential) reason to go skydiving and not be motivated to do so, if I believe that my reason is overridden. So internalism holds at most for overriding requirements. However, in the area of morality, I can believe that I am morally required to report my betting income to the IRS, but, if I am sure that I won't be caught (and that I won't feel guilty), I might have no motivation to report it. Even in the are of rationality, where internalism is more plausible, if I believe that it is irrational for me not to go to a psychiatrist, I still might

not be motivated to go, if I also believe that I am not worthy of being helped. So internalism fails in all of these areas.

Gibbard seems to reject this simple internalism.[13] If so, this is a virtue of his theory. But now he needs another way to specify what it is for a person to accept a piece of normative content. His proposal is a functionalist theory of normative avowal and normative governance:

Accepting a norm is whatever psychic state, if any, gives rise to this syndrome of avowal and governance by it. (75)[14]

Acceptance thus seems to be defined so as to give rise to each of two effects: normative governance and normative avowal.

Let's start with normative governance, since that is Gibbard's replacement for internalism. He defines it as follows:

Working out in community what to do, what to think, and how to feel in absent situations, if it has these biological functions, must presumably influence what we do, think, and feel when faced with like situations. I shall call this influence *normative governance*. (72)

What is new about normative governance is the source of motivation: normative discussion. When someone smokes for pleasure or out of habit, her motivation does not arise from normative discussion, so she need not accept norms that require or even permit smoking. In contrast, if we discuss famine relief and conclude that we ought to contribute, and this motivates us to contribute when the next famine occurs, then we are normatively governed.

If this is the point, Gibbard is still committed to the internalist claim that acceptance of norms implies motivation. (cf. 80, 100, 112–3) He has narrowed the relevant kind of motivation, but that only makes it harder to believe that this special kind of motivation must always accompany acceptance of any norm.[15] Even if I discuss tax evasion, and I conclude that I am morally required to report my betting income to the IRS, I still might have no motivation to report it. I also need not "feel" (72) any particular way when I reach this conclusion or when I omit this income on my tax form (whenever that is). The discussion did influence what I "think" (72), but there is nothing expressivist about that. Besides, Gibbard is supposed to be analyzing what I think, and it would be circular to say that a person believes a normative content when she is disposed after discussion to think it.

Gibbard might respond that this problem is solved by the other part of his theory: normative avowal. But this does not really help. Gibbard defined acceptance as what gives rise to a conjunction of "avowal *and* governance" (75, my emphasis), so acceptance still implies governance, and that is enough to create the problem.

Gibbard might make his theory disjunctive, so that acceptance requires *either* avowal *or* governance, but that leads to problems of its own. To see this, we need to look more closely at what Gibbard says about avowal:

To accept a norm, we might say, is in part to be disposed to avow it in unconstrained normative discussion, as a result of the workings of demands for consistency in the positions one takes in normative discussion. (74; cf. 234)

Normative discussion occurs "when we work out at a distance, in community, what to do or think or feel in a situation we are discussing." (73) Such normative discussion is unconstrained when the discussants "relax the psychological mechanisms of self-censorship" (74),[16] although they are still constrained by their "reciprocal" demands for consistency. (75)

Here Gibbard makes two separate claims. The first is that, if people accept a norm, then they will avow it at a distance when unconstrained. This seems OK. But he also claims the converse: if people avow something at a distance when unconstrained, then they accept it. This seems wrong. When I teach an introduction to philosophy, I constantly run into students who relax their self-censorship and avow Berkeley's idealism just in order to test the consistency of the position. But they don't really believe it. The same thing happens in normative discussions. Students often avow extreme libertarianism or nihilism just in order to see whether they can defend it consistently. But they don't really accept or believe it. Thus, distance and the absence of self-censorship do not always make people avow what they really accept and believe. Gibbard might respond by revising his account of "unconstrained normative discussion", but that will not be easy.

In any case, a more fundamental problem awaits. If Gibbard claims that avowal without governance is enough for acceptance, his theory ceases to be distinctly expressivist. The most descriptivist of Gibbard's opponents can agree that to believe a normative judgment is at least in

part to be disposed to avow it under certain circumstances. What is distinctive about normative belief, according to Gibbard (100), is its link to motivation and action. Without that link, Gibbard's theory reduces to a general formula with none of the implications that he claims.

Thus, Gibbard's theory of normative acceptance falls into a dilemma. If acceptance does imply normative governance, he cannot handle cases like reports to the I.R.S. If acceptance does not imply normative governance, he has surrendered the expressivist core of his theory.

5. OBJECTIVITY

Another area where Gibbard advances beyond traditional expressivism is in his analysis of objectivity. Unlike some expressivists, Gibbard realizes that:

> When a person calls something rational, he seems to be doing more than simply expressing his own acceptance of a system of norms. . . . He claims the backing of considerations that, in some sense, "compel acceptance" of what he is saying. Perhaps he is wrong, but that is the claim he is making. Any account of his language that ignores this claim must be defective. (153)

In short, "normative language does involve claims to objectivity in some sense." (154) In part III of his book, Gibbard attempts to define the relevant kind of objectivity.

The main problem arises from optional norms. I can accept a norm against drinking alcohol without believing that other people must accept it as well. Here Gibbard says that I have an existential commitment not to drink, and he admits that I do not really believe that drinking is irrational. So Gibbard needs to explain the difference between accepting a norm as an existential commitment and accepting it as a norm of rationality.

His solution lies in higher order norms:

> To accept a norm as a requirement of rationality, we might say, is to accept it along with higher order norms that require its acceptance. To treat it as an existential commitment is to accept it along with higher order norms that permit it, but that permit accepting at least one incompatible alternative. (169; cf. 171)

This is ingenious and useful. It does seem that, if I accept a norm as a requirement of rationality, I accept it as binding on everyone.

But the converse is not as obvious. I do not accept a norm as a requirement of rationality whenever I accept it as binding on everyone. Suppose I accept a first-order norm that forbids me to volunteer for any war, and I also accept a second-order norm that requires everyone else to accept this norm. I still might believe that my acceptance of this second-order norm is optional. More technically, I might accept both of these norms along with a third-order norm that permits them but also permits an incompatible alternative. Then I have an existential commitment to my second order norm. In this case, I would not believe that my first-order norm is a requirement of rationality. In order to believe that it really is irrational to volunteer for war, I must also believe that everyone is rationally required to accept not only the first order norm not to volunteer for war but also the second order norm that requires everyone to accept the first order norm.

Gibbard might respond by adding that rationality also requires a third order norm that requires everyone to accept my second order norm. But the same basic problem arises again at each higher stage. Why stop at the third order? What if I have only an existential commitment to my third order norms? Or fourth? Or fifth? It is hard to see why we should stop at any particular level. But if we do not stop, we never get a complete analysis of objectivity. I do not see how Gibbard can escape this infinite regress.[17]

This regress does not arise for Gibbard's opponents. If to believe that it is irrational to volunteer for war is to believe that this judgment corresponds to an independent reality, then this independent reality should rule out any norm of any order which permits anything contrary. There still might not be such normative facts. That is an ontological issue. But Gibbard admits that common normative beliefs and statements include an implicit claim to objectivity. (153—5) If so, this is an aspect of common thought and talk which his opponents can capture more naturally and without the infinite regress that Gibbard faces.

6. THE VARIETY OF NORMS

The final traditional problem for expressivism arises from the variety of norms. One thing that an analysis of a kind of judgment is supposed to do is to help us distinguish judgments of that kind from judgments of

other kinds. Gibbard accepts this burden when he tries to distinguish moral judgments from other kinds of judgments on the basis of his claim that moral judgments are about when anger and guilt make sense. But we also need to know how to distinguish judgments of rationality from other kinds of normative judgments.

His basic theory is that to call something rational is to express one's acceptance of a system of norms that permit it (or do not rule it out). This claim runs in two directions. He claims that, whenever someone calls something rational, the speaker expresses acceptance of norms that do not rule it out. He also claims the converse: whenever a speaker expresses acceptance of norms that do not rule out an act, the speaker calls the act rational.

The former claim might seem innocuous, but it does run into some problems. People often call a particular act rational or irrational without formulating any general norm (much less any system of norms) that permits or forbids that act. For example, I believe that it is rational to scratch my head right now, but I doubt I could formulate general norms about it. Even if I did, my confidence in the general norms would be nothing like my confidence in the particular judgment. This makes it odd to analyze my particular belief in terms of general norms.[18]

The problems for the converse are even more serious. It is easy to think of cases where a speaker expresses acceptance of norms that do not rule out an act, but the speaker does not call the act rational. If I say, "It is not impolite to smoke when you are alone," I express my acceptance of norms (of politeness) which do not rule out smoking alone. But I do not call this act rational, and I need not think that it is rational. And the same point could be made using norms of logic, language, morality, and so on.

Of course, this is not what Gibbard meant to say. Judgments about rationality do not express one's acceptance of just any norms. They express one's acceptance of norms of rationality. Since there are other kinds of norms that might conflict with norms of rationality, any analysis of rationality in terms of norms must specify the kind of norm in question. If this specification is made explicit, Gibbard's analysis would claim: A believes x is rational if and only if A accepts norms of rationality that permit x. But this revised analysis is circular on its face. Judgments of rationality are analyzed in terms of norms of rationality,

but this does not give us any clue about how to distinguish norms of rationality from other kinds of norms.

Gibbard might try to avoid this problem by saying that judgments of rationality are not about a special kind of norm but are about one's overall system of norms. Then his theory claims that: A believes x is rational if and only if A's overall system of norms permits x. But that still won't work. Suppose I accept norms of rationality that permit me to steal your car. I then believe that it is rational and not irrational for me to steal your car. I can also accept separate norms of morality that forbid me to steal your car. Then my total system of norms does not permit me to steal the car. But I still think that it is rational to steal the car. Consequently, beliefs about rationality cannot be analyzed in terms of the overall system of norms that I accept.[19]

The only solution seems to be to add a separate specification of the content of rationality. Gibbard does say some things about rationality early on, but they do not solve this problem. For example, he says that "rational" means "makes sense", (7) but this only explains the obscure by the more obscure. He also says that "rational" means "a flavorless recommendation on balance" (49), but to call it "flavorless" is no help in distinguishing judgments of rationality from other normative judgments. And he cannot define rationality by its effects on motivation and action, because then he just returns to the problems of internalism. So I must confess that I do not see any adequate way for Gibbard to distinguish judgments of rationality from other normative judgments.[20]

7. MORALITY AND EMOTIONS

Gibbard has similar problems distinguishing moral judgments from other kinds of judgments. On Gibbard's account, an act is morally wrong just when the agent would be to blame if he were responsible, (44) so the basic issue is when an agent is "to blame". Gibbard then claims that "To say that he would be to blame is to say that it would be rational for him to feel guilty and for others to resent him" (45; cf. 126) or "to feel angry at him" (44; cf. 150). Gibbard later admits that this is not quite right, because morality is impartial, but anger is not. (126—7) Nonetheless, he continues to use the terms "anger" and "guilt".

The real problem is to identify these emotions. If we cannot tell

when an emotion counts as anger or guilt, we cannot tell when a judg-
ment is a moral judgment. (cf. 272) But this is not easy. We cannot
identify anger or guilt phenomenologically or physiologically, because
they feel different and have different physical causes and effects in
different people and at different times.

One common approach is to identify anger and guilt by the moral
judgments that cause them, so that to feel guilty is to feel bad as a result
of one's belief that one did something morally wrong. Gibbard rejects
such "judgmentalism" on the grounds that "I can feel angry at you and
yet think that it makes no sense to do so; I can think that really you
have acted as you should." (130; cf. 148—9) Even if judgmental
theories can be saved from this criticism,[21] Gibbard still cannot adopt a
judgmental theory. As he says, if emotions cannot be identified inde-
pendently of moral judgments, "the norm-expressivist account of moral
judgments is circular: it invokes anger and guilt to characterize moral
judgments, but we must understand moral judgments already if we are
to characterize guilt and anger." (148; cf. 128) So Gibbard needs some
non-judgmental way to identify anger and guilt.

Gibbard discusses several possibilities[22] but eventually settles on an
attributional theory. (143) On this theory,

When I think of myself as guilty, I see myself as being in a state that I conceive as
follows: it is typically caused by my own acts of certain kinds, it is expressed by a guilty
mien, and it typically moves me toward apology and amends. (148)

Of course, my guilt is not caused by my act of a certain kind unless I
believe that I did an act of that kind. This belief might seem to reintro-
duce circularity, but it does so only if the relevant kinds are defined by
moral judgments. In order to avoid this circularity, Gibbard claims that
an observer who wants to characterize guilt

can appeal directly to the central kind of circumstances in which, in my culture, one is
thought to be at fault. Those circumstances will be part of the cluster the observer uses
to define guilt. (149)

Guilt is thus defined by paradigm examples in one's culture and then
extended to other cases that are similar in some way.

This runs into several problems. The first is that moral rebels and
reformers might reject the paradigms of their cultures. For example,

people who are vegetarians for moral reasons usually feel bad after they eat meat, but are these feelings "guilt"? Vegetarians certainly realize that eating meat is not a paradigm of fault in our culture. They do not feel much like they feel when they lie or break a promise. They do not typically display "a guilty mien". And they are not typically moved "toward apology and amends." (Who would they apologize to?) So their emotions do not have any of the traits in Gibbard's account of guilt. But they still feel guilty. At least they usually describe their feelings that way. Why? The reason seems to be that they believe that they did something wrong. Without reference to these moral judgments, it is hard to see how Gibbard could recognize such feelings as guilt. But with moral judgments, his theory lapses into circularity.

A different circle arises when we ask how to determine which are the central examples of a culture. If we do not yet know which emotions count as guilt, and we also do not know which judgments count as moral judgments, how can I identify the central examples where "in my culture, one is thought to be at fault" morally? For example, where I was brought up, it was considered wrong to put the fork on the right side of the plate. This was one paradigm case of a faux-pas. But it would be at best misleading to use this case to define guilt and moral judgments. Why? Because this mistake is one of etiquette, not of morality. But to see this requires a moral judgment: that it is not *morally* wrong to put the fork on the right. So we have to assume this moral judgment even in order to pick out the paradigm cases. Thus, Gibbard's theory eventually relies on moral judgments to pick out moral judgments, and he falls back into the kind of circularity that he was trying to avoid.

8. CONCLUSION

I conclude that Gibbard fails to solve several of the traditional problems for expressivism. He solves some of these problems, but his solutions to them in effect give up expressivism. Of course, one might respond that it does not really matter whether his theory is expressivist. In some ways, I agree. Gibbard says many fascinating things about morality which have at most indirect connections to his expressivist analysis. I am thinking especially of his later discussions of hyper-

scepticism (180), parochialism (203 ff.), and indirect pragmatism (224). These views could still be developed even if he gave up expressivism. All I have tried to show here is that he does need to give up expressivism unless he can solve the problems that I have raised.[23]

NOTES

[1] Cambridge; Harvard University Press, 1990. All references in the text and notes are to this book unless otherwise indicated.

[2] Of course, everyday talk is vague and varied, so Gibbard sometimes seems to admit that his analysis does not quite fit some aspects of ordinary thought and talk (49, 124n, 154—5, 175, 284). Nonetheless, Gibbard is not a stipulator or a reformer, and he does not want to analyze only technical or philosophical talk.

[3] Sometimes Gibbard states his view as the claim that "normative judgments are not *pure* judgments of fact" (105, my emphasis; but contrast the next sentence). In these qualifications, he seems to be admitting only that normative judgments can state non-normative facts. He still denies that normative judgments state normative facts. That denial is what I call his negative claim.

[4] Negators include prescriptivists if prescribing is different from expressing acceptance of a norm. Gibbard seems not to distinguish these speech acts, but one can prescribe without expressing acceptance (e.g. if Ann says, "Watch your head", solely because her boss told her to say this to people in a doorway), and one can also express acceptance of a norm without prescribing (e.g. if I say, "I think that people ought to use chopsticks in Chinese restaurants, but you do what you want.").

[5] Gibbard takes his main opponents to include normative realists, including Sturgeon, Railton, and Boyd (34n, 107, 116n, 122) and ideal observer theorists, such as Brandt (18—22, 183—8). These opponents deny Gibbard's negative claim, but they do not *have to* deny his positive claim (or his naturalism). So Gibbard needs to argue not only against the purely descriptivist versions of these positions but also against the mixed versions.

[6] Gibbard calls these combinations "factual-normative worlds" (95), but this is misleading. If there are no normative facts, as he claims, then the "world" is no different when the same facts are combined with different norms. This is impossible if normative claims supervene on non-normative ones, but Gibbard gives no reason to believe in the supervenience of the normative. This is a problem for Gibbard, because, if supervenience is a feature of common normative language, as many think, Gibbard needs to explain it.

[7] A normative system is defined (87) as a combination of norms that forbid, require, and permit, plus rankings of these norms. Gibbard says norms are imperatives (70) and often states norms in the imperative. (e.g. 91, 98, 165, 168) But he also crucially says that norms plus facts entail normative judgments (95), which are indicative, and he never tells us how imperatives entail indicatives.

[8] See Simon Blackburn, *Spreading the Word* (Oxford; Clarendon Press, 1984), pp. 189—96; and my review in *Philosophy and Phenomenological Research* 48 (1987), pp. 163—6. Similar problems arise for other embedded contexts which are not assertive: I do not seem to express acceptance of any norm when I utter questions ("Is murder wrong?"), disjunctions ("Either euthanasia is wrong or abortion is not"), or negations ("Sleeping is not irrational"). Gibbard also needs to explain what is going on in these contexts.

[9] This is a Gibbardian way of putting Davidson's theory in "Causal Relations", *Essays*

on Actions and Events (New York; Oxford, 1980). Just as Davidson claims that to say "a causes b" is to claim that there is some causal law linking a and b without committing oneself to any particular causal law, so Gibbard claims that to say "a is rational" is to express acceptance of some norm that permits it without committing oneself to any particular norm. This is how Gibbard solves the problems of communication and naïveté. (97)

[10] Gibbard does use the term "express", but a speaker can express beliefs as well as emotions and desires, so even a realist can agree that to call something rational is to express a state of mind, namely, a belief. The debate is not about whether statements express states of mind. It is about what kinds of states of mind are being expressed. Gibbard describes these states of mind as acceptance and ruling out, but even when I express a factual belief that P, I express "acceptance" of P, and I "rule out" what is incompatible with P. So the difference between Gibbard and his opponents must lie elsewhere.

[11] Gibbard also claims that inconsistency leads to "a special kind of self-frustration" which we cannot "live with" (289). But I might not be frustrated at all if my inconsistent beliefs are popular enough and do not affect my life too much.

[12] Internalism about reasons is less strange, and a confusion of these two claims might explain the popularity of internalism about motives. See David Brink, *Moral Realism and the Foundations of Ethics* (New York; Cambridge, 1989), chapter 3.

[13] Gibbard never explicitly denies simple internalism, but he implies as much when he distinguishes internalization from acceptance (71) on the basis of a more complex example: Milgram's experiments (58—60). According to Gibbard, a subject in these experiments "thinks that it makes no sense to cooperate" (60) because he "accepts as having most weight in his situation . . . norms that turn out to prohibit cooperation." (61; cf. 71) Nonetheless, the subjects still consistently act contrary to what they think and accept. Thus, as Gibbard presents it, this example refutes simple internalism.

[14] One could interpret this definition so that it claims only that acceptance is the state that causes avowal and governance when these effects do occur but it allows that acceptance can still exist without causing any avowal or any governance. But if this is all Gibbard claims, he has not told us what makes this the same state (acceptance) when it does not cause either avowal or governance, or how we can tell whether someone does accept a norm if he neither avows nor is governed by it. This reading also leaves it unclear how acceptance is any different from normative beliefs on descriptivist views, since these opponents can agree that normative beliefs *usually* cause avowal and motivation. For these reasons, I will focus in the text on the conjunctive and disjunctive interpretations.

[15] An additional problem is that I cannot be motivated in Gibbard's special way unless my motivation arises from normative discussion, but I can accept norms that I have never discussed at all. In fact, Gibbard says that some norms are not linguistic, (69—70, but contrast 46, 57) so they could not be discussed. This also raises problems for what he says about avowal below, and for some of his evolutionary explanations.

[16] This is the only constraint Gibbard mentions. Presumably, other constraints (e.g. time) are also lacking, but no list is provided.

[17] Gibbard tells us (169 note 12) that his use of higher orders was influenced by Harry Frankfurt, "Freedom of the Will and the Concept of a Person". A similar regress in Frankfurt's theory was pointed out by Gary Watson, "Free Agency" (sec. III). Both articles are reprinted in *Moral Responsibility*, ed. John Fischer (Ithaca; Cornell, 1986).

[18] Related points about appeals to rules in morality are made by Jonathan Dancy, "Ethical Particularism and Morally Relevant Properties", *Mind* 92 (1983). If particularist claims even make sense, they pose problems for analyses of normative judgments in terms of general norms.

[19] On Gibbard's analysis of morality (below), the overall system of norms says that it is rational to steal the car and also rational to feel guilty afterwards. This might seem to

permit me to steal the car, as long as I feel guilty afterwards. However, this only shows that Gibbard has failed to capture the force of moral norms. Moral norms are not like rules in parking lots where one is permitted to park if one pays the fee, or even like parking laws, since some see parking fines as mere fees. Anyone who really believes that stealing is morally wrong does not believe that one is permitted to steal as long as one pays the price in guilt and in jail.

[20] This problem is compounded by a basic conflict in Gibbard's theory of rationality. Gibbard usually defines "rational" in terms of what norms "permit" (7, 46, 83, etc.), and once he says explicitly that "the term 'rational' is permissive rather than obliging." (89) This fits his identification of what is rational with what "makes sense" (6—7, 89). In other places, however, he says that "to think something rational is to accept norms that prescribe it" (47; cf. 70) and to call something rational is to "recommend" it (49). He also says that he reserves "rational" and "makes sense" for what is "well founded or warranted" (37; cf. 187n) and that judgments that an act is rational "settles what to do" (49). But he can't have it both ways. In common use, norms that permit skydiving or say it makes sense do not prescribe or recommend skydiving or say it is "well founded or warranted". Gibbard tries to solve this problem by substituting "makes the most sense" for merely "makes sense" (7n), but then he admits that this is inadequate because it "carries an unwanted claim of uniqueness" (7n). So it remains unclear whether he uses "rational" as a permissive term.

[21] One way to save judgmental theories is to insist only on a judgment of prima facie wrongness at some time. My emotion does not seem to be guilt if I never believed that the relevant kind of act was even prima facie wrong. Another response is to require not full beliefs but only construals, appearances, or other belief-like states directed at moral judgments. Cf. R. C. Roberts, "What an Emotion Is; A Sketch", *Philosophical Review* 97 (1988); and Patricia Greenspan, *Emotions and Reasons; An Enquiry into Emotional Justification* (New York; Routledge, Chapman and Hall, 1988). These accounts of emotions would still make Gibbard's theory circular, for he would have to analyze these weak moral judgments and belief-like states in terms of guilt and anger.

[22] Gibbard also discusses an adaptive syndrome view, which defines anger and guilt by their biological functions (134; cf. 147), but he explicitly concludes, "The explanation of what these emotions are cannot be the one I have sketched." (141) He bases this rejection on anthropological data. (140—1) I would add that different emotions could have the same evolutionary function, and the same emotion could have different evolutionary functions. Furthermore, it is only general kinds of emotions that have evolutionary functions. This makes it hard to see how to use such functions to determine whether an emotion in a particular case is anger or guilt. But that is what is necessary to tell whether a particular judgment is a moral judgment. There are also problems for Gibbard's claims about the specific functions of these emotions. He claims that guilt "aims to placate anger" and thereby to avoid social friction, and that anger is aimed at insufficient or bad motivation and leads to punishment. (138—9) However, I can be angry at a business partner who makes a useless purchase of goods, even if he had no bad motivation and does not deserve any punishment. And I can also feel guilty even when I think nobody will be angry or punish me.

Department of Philosophy
Dartmouth College
Hanover, NH 03755-3592
USA