ALLAN GIBBARD

REPLY TO SINNOTT-ARMSTRONG

I am grateful for Walter Sinnott-Armstrong's careful and extensive commentary on my book. His commentary is rich and wide-ranging, so I must confine myself to just some of the points he raises.

Let me start with a word about meanings and normative facts. Sinnott-Armstrong distinguishes what there really is, on the one hand, from what we believe and mean. I could hardly reject the distinction: It indeed is one thing to ask whether there really are normative facts, and another to ask whether we believe that there are — and whether our normative statements mean that certain normative facts obtain. Still, I think we should be charitable enough with ourselves to see the two questions as related. It's a chief aim of my book to argue that if we didn't believe in normative facts, we could coherently go on talking and thinking, normatively, pretty much as we do talk and think. If, then, we convince ourselves that the world doesn't contain normative facts, why accuse ourselves of having been in error? If we've been talking pretty much as we should in a world without normative facts, and thinking pretty much as we should in a world without normative facts, why not pat ourselves on the back and say we've been talking and thinking pretty much correctly?

I'm not, then, proposing a blanket reform in our normative thought and talk. I am proposing that some of us philosophers should change what we say about normative thought and talk. I'm proposing that there are limits to the analogy between *ought*-statements and statements that less problematically aspire to state facts. I'm proposing that the best way to theorize about normative thought and talk is not to revel in this analogy, but to look more closely at normative psychology and sociology, and draw on logical analogies with imperatives.

1. EXPRESSIVISM AND MIXED THEORIES

My own view of normative language is expressivistic. Normative state-
ments express states of mind — "normative judgments", we can call
them — and the way to explain what normative statements mean is to
say what normative judgments are like. They are not, I claim, best
explained as states of belief in facts that are somehow normative; they
are best treated in another way.

Sinnott-Armstrong asks why I don't go for a mixed theory of norma-
tive language, partly descriptivistic, partly expressivistic. Why don't I
think that normative statements do two things at once: state facts, and
also express states of mind that aren't factual beliefs. A part of the
answer is just that I'm not sure how such a mixed theory is supposed to
work. Some normative terms do clearly involve an interplay of descrip-
tive and non-descriptive factors. In a recent paper I take up the word
'lewd' as an example (Gibbard, 1992b). Its meaning combines descriptive
elements — the notion of sexual display — with normative elements that
I interpret as non-descriptive. But this interplay, I find, is none too
simple, and in the case of general moral terms like 'wrong', I don't seem
to need it. An expressivistic analysis seems to do the job by itself.[1]

What would the descriptive part of a mixed theory be supposed to
accomplish? My opponents, Sinnott-Armstrong says, are driven to
include descriptivist elements in their analyses by the need to explain
such things as modus ponens. I say one doesn't need descriptivist
elements to do this, and if one did, they wouldn't manage the job. Think
of emotive discourse: From "Yea for all Greeks" and "Socrates is a
Greek" follows "Yea for Socrates". To say this, we don't have to
suppose that "Yea for" has a descriptive meaning. And with mixed
emotive-descriptive language, validity doesn't seem to be settled by the
validity of arguments couched in descriptive language. We can see this
with racial epithets that we, as decent people, would reject. Take
"nigger": The term seems to mix description and emotional expression.
Suppose, then, a racist reasons this way: Premise: "King had dark-
skinned African ancestry." First consequence: "King was a nigger!" And
let the next inference be just to "King!" said in a tone expressing hatred.
The initial premise seems uncontroversial, but then I want to object to
the reasoning at the first inference, where the pejorative expression

kicks in. Why do I find the statement "King was a nigger" totally unacceptable? Partly, because accepting it would seem to commit me to hating King on account of his race — and I reject and abhor such hatred. If, though, I thought that logic follows the descriptive component, I'd think that in accepting the premise "King had dark-skinned African ancestry", I was committed to accepting the first inference in this chain of reasoning, and so accepting "King was a nigger". Things don't seem to work this way.

Still, as Sinnott-Armstrong points out, normative logic can be much more complicated than the emotive logic I've just been using as a test. I myself end up with a rather complex account of normative logic, and Sinnott-Armstrong asks whether the theory that results is still in any way expressivistic. Fair question! Of course the normative logic itself isn't distinctively expressivistic, since I was trying to show that on a theory like mine, ordinary logic would apply. But in a real sense too, the theory, once it has undergone this transformation, is no longer expressivistic. Neither, though, is it descriptivistic. The meaning of a normative statement, I say, is a matter of the combinations of norms and descriptions the statement rules out. I don't, then, directly characterize all the states of mind that can be expressed by normative statements. Rather, many of them I characterize indirectly, by the inferential relations they stand in. Some of the states in this inferential web are states of what I call normative governance: a special kind of tendency to do (or believe or feel) what a normative dictate tells you to do (or believe or feel). It is this special inferential tie to normative governance, I claim, that distinguishes normative statements from purely descriptive statements.

I could call my position, then, a "loose" expressivism rather than a strict one. A *strict* expressivist for normative language would provide a more or less direct characterization of a state of mind all such language expresses: the state of mind isn't a belief, but an emotion, say, or a state of accepting norms. As a *loose* expressivist, I take such a state of mind as explanatorily basic. In this, I'm unlike a descriptivist, who takes something quite different as explanatorily basic — namely facts (or putative facts) that somehow, if they obtained, would constitute normative facts — and who then explains normative judgments as beliefs in these normative facts. What I myself take as explanatorily basic is a

state that goes with normative governance. We might call this state accepting a *normative directive*: a normative requirement that applies to oneself right now. (At points in the book, I make it sound as if the explanatorily basic state is accepting a system of norms, but I see now that I should have treated normative directives as basic.) Only a limited range of normative language, though, expresses the speaker's accept-ance of normative directives. The states of mind expressed by norma-tive statements outside this limited range are to be explained by their inferential relations to this explanatorily basic kind of state.

What is distinctive, then, about all expressivism, strict or loose, is an order of explanation: one begins with a special kind of psychic state which is not to be explained as straight belief in some peculiarly normative kind of fact. In my own case, I begin with the state of mind: accepting a normative directive that applies to oneself right now. And I explain the states of mind expressed by a wider class normative state-ments by their inferential relations to this explanatorily basic kind of state of mind.[2]

2. NORMATIVE INTERNALISM AND DESCRIPTIVISTIC THEORIES

That brings me to the set of issues that go under the name "inter-nalism". We both agree that normative convictions are closely tied to motivation. But this tie might be like the one between believing that it's raining and being motivated to carry an umbrella. I myself claim that the tie is of a different kind: that it is a matter of meaning. Barring special states — despondency, say, or exhaustion — you can't be said to accept that you rationally must do something without having some motivation to do it. This motivational tendency might be outweighed by enticements to the contrary; weakness of will provides no counter-example to *this* version of internalism.[3]

What, then, of complete but unprincipled indifference to rationality? Well, as I think about it now, I do want to admit that in depressed or abnormal states of mind, the dictates of normative governance might be rendered totally ineffective motivationally. That's still consistent with the rest of what I want to say: that normative judgments are best characterized by their logical ties to a state that is normally motivating

— to a state of having decided what to do or what to believe or how to feel, in so far as discussion and reasoning can decide such things.

Still, in normal states of mind, I'm saying, I don't sincerely believe that a kind of act is wrong if I have no tendency to feel guilty upon performing acts of that kind. This tendency may be outweighed, but it has to be there. Sinnott-Armstrong asks, then, can't I accept that it's wrong, say, not to report betting income on my tax return, and still have no tendency at all to feel guilty when I don't?

I'm not convinced I can. Suppose neither of us reports our betting income, and neither of us has the least tendency to feel guilty for not doing so. (Perhaps we're both on principle indifferent to moral considerations.) You, though, think that we're acting wrongly, morally, in not reporting our betting income, whereas I don't. You think it is morally wrong to disobey the law, whether or not anyone else obeys it. I think it would be morally wrong to disobey the law only if we thereby became free riders on others' moral restraint. We both agree on what the income tax laws say, and we both agree that virtually no one reports informal gambling earnings. So in what do we disagree? Is it just in the application of a word? You have no tendency at all to feel guilty for not reporting your betting income. You have no tendency to resent others who don't report betting income. What's at stake in whether you convince me? Aren't we just going through the forms of a moral dispute, using the word 'wrong' with no meaning?

We can say that we take a great theoretical interest in moral right and wrong, even though deciding an act is wrong or not makes no difference to how we think it makes sense to feel about the act. But isn't this a theory without a real subject matter? We can say that we're trying to explain our patterns of judgment in a principled way. But our judgments seem to take different patterns, and perhaps either of us can find principles to match his judgments, if we bother to play that game. Are we therefore just signifying different things by the word 'wrong', so that our disagreement is only apparent? What would make this moral disagreement genuine, it seems to me, is the normal tie to sentiments of guilt and resentment. Then something would be at stake, namely how to feel about the non-reporting of betting income.

Of course I can adopt a principled indifference to the kinds of considerations other people treat as moral — without drawing too fine a

line about what those considerations are, and how they legitimately
weigh in judgments of moral rightness and wrongness. What I can't do,
in that case, is to enter into moral disputes the way other people can. If
I use the term 'morally wrong' to apply to things most people regard as
wrong, I'm then not using the term 'wrong' to mean what they do. I
can't then have meaningful, fine-grained moral convictions on matters
of moral controversy.[4]

Still, with English terms like 'wrong' or 'irrational', meanings are
bound to be controversial. Whether 'ought'-judgments and the like are
tied to normative governance by the meaning of the English term
'ought' won't be entirely clear as a matter of actual English usage. The
more important question, though, is not how our language currently
works, but how to sharpen it. Shall we purge it of all non-descriptive
elements and work to make sure that all our words have agreed truth
conditions? Or shall we cultivate language also that has an analytic tie
to normative governance? Shall we discuss and reason only about how
things are factually? Or shall we cultivate the linguistic resources to
discuss also how to act and think and feel in a wide variety of
situations?

With the issue put this way, I myself will extol the advantages of the
latter kind of reasoning and discussion. I'll sketch how we can think of
past, future, hypothetical, and even fictional cases, and decide together,
hypothetically, what to do or to think or to feel in such cases. I'll show
how this constitutes a kind of rehearsal for situations that may come up
as we live our lives together. I'll stress the advantages of not restricting
in advance the kinds of considerations that can be put forth in such
discussion. Descriptivistic opponents will claim for their side the
advantages of getting all meanings tied down to facts. I'll say I welcome
having meanings tied down to facts, and indeed I can welcome all the
clear, descriptive meanings they can invent. Still, I'll say, we need my
other kinds of meanings too — and if any meanings distinctively merit
being called normative, isn't it meanings analytically tied to normative
governance in this way?

3. NORMS OF RATIONALITY AND OTHER NORMS

Let me turn now to the question of what distinguishes various different
kinds of norms. I myself say that all norms are, at base, norms of

rationality. So what makes a norm a norm of morality, say, or a norm of politeness? Different kinds of norms, I answer, concern the rationality of different kinds of things. Moral norms, for instance, concern the rationality not of actions directly, but of moral sentiments toward actions.

Now Sinnott-Armstrong thinks that this won't work for such special-purpose norms as norms of politeness. What, on my theory, would it mean to say such things as, "It's not impolite to smoke in private"? The person who says this, I claim, expresses his acceptance or norms, but not norms governing private smoking directly. He expresses his accept-ance of norms governing some kind of reaction to smoking in private. So it's perfectly consistent to say, "It's not impolite to smoke in private, but it *is* irrational." This expresses one's acceptance of norms that forbid smoking, but also forbid certain emotional reactions to smoking — reactions, we might say informally, that constitute experiencing private smoking as rude.

But what are these reactions? I should say first, I'm talking about someone who is using talk of politeness normatively. As with morality, a person can use seemingly normative terms sociologically, describing convictions by mouthing them — just as a historian who wants to describe early modern witchcraft beliefs without buying into them can say things like "Witches had sexual intercourse with the devil." In a like vein I might say, "It isn't polite for a woman to reveal herself as brighter than the man she's with." If I said this, I assure you, I'd be speaking at most sociologically, nor normatively!

What am I doing, then, when I speak normatively in terms of polite-ness — when I buy into standards of politeness? (Of course as Philippa Foot says, I'll still want to allow that at times, a little rudeness is in place.[5]) To call behavior impolite is to express norms governing a reaction to it, I'm saying, but *what* reaction? Well, I don't imagine we'll get a terribly precise analysis of what the term 'impolite' means, whether my way or anyone else's. To get in the right ballpark, though, we'd have to be more sensitive to the role of politeness — in the sociological sense — than has been good intellectual form in the last couple of centuries.[6] We'd have to be pretty Goffman-like,[7] or take Confucian treatments of etiquette seriously.[8] We'd have to recognize how much of the processing power of our brains is devoted to our standing without other people. We'd have to think about the roles that

emotions of respect and disrespect, offense, humiliation, and the like play in human life.

As a first stab, I'd say this: When we discuss politeness and impoliteness normatively, we're working out when to be offended and when not to be. (I think the feelings associated with normal, polite conversation amount to more than lack of offense, but we don't seem to have a name for our reaction to normal social reassurance and affirmation.) Perhaps the best thing to say, then, is that impolite acts are acts that *prima facie* warrant feelings of offense. I say '*prima facie*' here, because as with wrongness and blameworthiness, there can be excuses: the ignorance of a child or a foreigner, or provocation.

Suppose Emily says that it's irrational to smoke at all, but it's not impolite to smoke if you smoke in private. What is she doing, according to my theory? She's expressing her acceptance of norms that say not to smoke, but also not to be offended by private smoking. And isn't this pretty much what she *would* be doing if she said this?

4. MORAL SENTIMENTS

I characterize various different normative concepts, then, in terms of the different feelings that pertain to them. Now a frequent way of criticizing this way of doing things is to claim that we can't identify an emotion except in terms of beliefs. We can't understand, say, feeling offended without first having a descriptive characterization of impoliteness. Sinnott-Armstrong says, "We cannot identify anger or guilt phenomenologically or physiologically, because they feel different and have different physical causes and effects in different people at different times." If things were that bad, though, I'm baffled how we would ever have learned to speak about emotions. I think we're capable of a lot more self-monitoring than the mainstream of current anti-Humean emotion theory will allow. I don't now have to say "I did wrong, so this feeling must be guilt" — although as with pain, to be sure, various external cues must have played a role in my learning term in the first place. Suppose, though, I have cause both for guilt and for fear, and I have a bad feeling. I may be able to sort out whether this feels like guilt or like fear without establishing the etiology of my feeling, and without finding a proposition embedded in it.

Still, the theory of emotions is difficult and problematic. In the book I put forth what I called an "adaptive syndrome" view of emotions, but I also confronted some anthropological reports that seemed hard to square with this view. Anger pretty clearly has cross-cultural patterns and a continuity with animal emotions, but I was much more puzzled about guilt, and the distinction between guilt and shame.[9]

I very much want to resist any assumption that a psychic state must be a factual belief or else fit for the philosophical garbage pail. As time goes on we're getting more and more evidence that a big chunk of specifically human brain tissue ties social cues to bodily feelings, and that damage to these areas can cripple a person for social, human living, without degrading such things as performance on IQ tests.[10] And so I'm finding myself better confirmed now than when I wrote the book in my speculation that places moral judgment among these intricate workings.

I'd expect, then, that by evolutionary design, different felt bodily reactions will go with different mechanisms for dealing with social cues, and for prompting the kinds of motivation that, with our ancestors, tended to promote reproduction when such cues were received. That won't mean that these mechanisms are currently adaptive, it should be stressed, especially in the strict biological sense. (We know, for instance, that educated people in rich industrial societies have a low rate of what evolutionary biologists call "reproductive success".) But there will be lots of specific emotional mechanisms that were shaped intricately by the Darwinian selection pressures of proto-human social life. And there's no reason to think that these mechanisms either fit neatly a cognitivist's picture of knower and facts known, or can be pushed outside the scope of human discussion and reasoning.

Of course even if this rough picture of human emotions is true as far as it goes, it leaves big psychological and philosophical questions: Where, if anywhere, does moral judgment appear in such a picture, and where do other kinds of normative judgment appear? What are the roles of social learning in the workings of these mechanisms? Clearly discussion and linguistically encoded thought have important roles, and I think that coordination — in Thomas Schelling's very broad sense — will be a crucial part of the story.[11] But are adaptive emotional mechanisms one by one subject to the kind of control that I've been

calling normative governance? Social learning will have important
effects on the workings of these mechanisms, and I've schematized this
learning under two broad headings: what I call "internalizing" norms
when one is in situations in which the norms apply, and what I call
"accepting" norms, as a result of normative discussion and internal
rehearsal for normative discussion.[12] The real story, though, will
doubtless be far more complex than this. I don't put forth my speculations
about feelings and accepting norms, then, as something we should
accept or expect later to accept. I do think, though, that my schema
makes a better starting point for further investigation than the belief-
desire explanations that philosophers currently find so appealing.

5. OBJECTIVITY

I spend a long time in the book asking what kinds of objectivity an
expressivist like me can claim for his normative judgments. Suppose I
accept the norm "Value accomplishment for its own sake." What distin-
guishes my thinking this a demand of rationality from my regarding it as
an existential commitment? I said the distinction is a matter of the
higher order norms one accepts. I might accept higher order norms that
permit me to reject this norm, and in that case, I regard valuing
accomplishment as an existential commitment of mine. If, on the other
hand, I accept higher order norms that tells everyone, mandatorily, to
accept this norm, then I think valuing accomplishment to be a demand
of rationality.

Now even if my second order norms are universal and mandatory in
form, Sinnott-Armstrong objects, I might accept third order norms that
permitted rejecting my second order norms — and then I wouldn't
regard valuing accomplishment as an objective demand of rationality.
Don't I, then, get myself into an infinite regress of higher and higher
order norms?

The criticism is a good one; it demands rethinking. What I now want
to venture is this: We don't ordinarily come to accept norms of any very
high order at all. Mostly we just don't think in terms of higher order
norms; the question of the status of the norms we accept doesn't come
up. Now if a person accepts only base order norms, it seems fair to say
that he hasn't formed any explicit opinion about the objectivity of any

of his normative judgments. If he unreflectively demands their accept-
ance on the part of everyone around, then he does *treat* the norms he
accepts as objective, but he does so unreflectively. Take next the case
where he explicitly accepts second order norms that require everyone
to accept his ground order norms. Suppose he treats these second-
order norms as objective, and does so unreflectively. Then shouldn't we
say that he explicitly thinks his ground order norms to be objective?

The default option, in effect, is to say that one unreflectively treats
one's norms as objective. Treating something as an existential commit-
ment, after all, is a highly self-conscious thing to do. Take, then, the
highest order norms a person accepts. If he is disposed unreflectively to
make conversational demands on others on behalf of these norms, then
he unreflectively treats them as objective. Proceed downwards in a
chain, then, from these highest order norms he accepts, through a set of
norms he accepts at each level down to the second. Suppose that the
norms in the set at each of these levels require *everyone* to accept the
norms in the set at the level below. Then he explicitly regards all the
norms in the sets below the top level as objective. If instead he accepts
a norm but there is no such chain leading down to it, then he regards
the norm as in some way non-objective — perhaps in quite a recondite
way.

The regress, then, stops at the highest level of norms that he accepts.
These he does not explicitly think objective or non-objective, though he
may treat them unreflectively as objective. Below that level, he may
explicitly regard various norms he accepts as objective, or he may
explicitly regard them as in some way non-objective. Whether he
regards a norm as objective depends on whether the right kind of finite
chain descends to it from higher order norms he treats as objective
unreflectively.

All this, I admit, is overelaborate as a description of how we actually
think. But it does show that no logical paradox lies hidden in the
possibility of thinking various norms one accepts to be objective or not.
One could, then, come to have explicit, consistent convictions on the
matter.

This and many of Sinnott-Armstrong's criticisms bear on the broad
question, What does psychological reality have to do with content and
with logic? This question raises a whole complex of philosophical

puzzles, puzzles that are extraordinarily vexing. The question applies to strictly factual beliefs, and it applies to whatever other kinds of psychic states may mimic factual belief is various ways. Indeed this broad puzzle, we might say, makes for *the* central cluster of problems that current philosophy is tackling. I can't hope to have gotten far with these problems, either in my book or here.

I do insist, though, that we shouldn't apply a double standard, talking as if these matters were unproblematic in the case of factual beliefs, and so problematic only for quasi-beliefs, as we might say: psychic states that mimic factual beliefs but aren't factual beliefs in the fullest sense. To call for a single standard, though, is not to solve these this cluster of problems. We philosophers have work to do, and incisive criticisms like Sinnott-Armstrong's advance the work greatly. I hope I've been able to suggest, though, at various points, why I might want to go along different paths from the ones his commentary seems to suggest.[13]

NOTES

[1] I argue this at length in Gibbard (1992a).
[2] If, as I do, we interpret R. M. Hare as an expressivist (see Gibbard, 1988), then, we can say, the classic statement of loose expressivism is Hare (1970). See also Blackburn (1984, 189—196) and (1988).
[3] Nor do some of Sinnott-Armstrong's other purported counterexamples work against this particular internalist claim: The claim isn't that beliefs in rational permissibility must motivate. And it certainly isn't that the motivations of normative governance must prevail.
[4] See Gibbard (1992a) for further discussion along this line.
[5] Foot (1958, 102—105).
[6] See Trilling (1972) for a fine discussion of this.
[7] Goffman (1959).
[8] I am grateful to Sin-Yee Chan for giving me a set of fascinating Confucian quotations about ritual and related matters. The subject looks extremely rich, and I wish I knew more about it.
[9] The book reads at one point as if I end up just rejecting the adaptive syndrome view, but this is misleading. "If," I say, "emotional conceptions differ as much as" certain "reports suggest, that seems to spoil the view of guilt and shame as distinct biological adaptations. The explanation of what these emotions are cannot be the one I have sketched" (141). I did regard the initial *if* as genuine, and indeed I'd regard it as a bigger *if* now than when I wrote the passage. Note too that my doubts were about guilt and shame; I can't find serious reasons for doubting that anger is a matter of an adaptive syndrome. (The psychologists Schacter and Singer (1962) adopt a theory that may be hard to square with this, and seem to talk as if the experiments they report demand a theory like theirs. I myself find the connection of their experiments to their theory quite tenuous.)
[10] See esp. Saver and Damasio (1991).

[11] Schelling (1960, Chap. 2).
[12] Both states are motivational, on my telling, but only *accepting* a norm involves "normative governance".
[13] This work was supported by a Guggenheim Fellowship, for which I am very grateful.

REFERENCES

Blackburn, Simon (1984). *Spreading the Word* (Oxford: Oxford University Press).
Blackburn, Simon (1988). "Attitudes and Contents". *Ethics* 98, 501—517.
Foot, Philippa (1958). "Moral Arguments". Reprinted in her *Virtues and Vices* (Los Angeles: University of California Press, 1978).
Gibbard, Allan (1988). "Hare's Analysis of 'Ought' and its Implications". D. Seanor and N. Fotion (eds.), *Hare and Critics* (Oxford: Oxford University Press, 1988).
Gibbard (1990). *Wise Choices, Apt Feelings: A Theory of Normative Judgment* (Cambridge, MA: Harvard University Press).
Gibbard, Allan (1992a). "Moral Concepts: Substance and Sentiment". James Tomberlin, ed., *Philosophical Perspectives* 6.
Gibbard, Allan (1992b). "Thick Concepts and Warrant for Feelings". *Aristotelian Society Supplementary Volume* 66, 267–83.
Goffman, Erving (1959). *The Presentation of Self in Everyday Life* (Garden City, NY: Doubleday).
Hare, R. M. (1970). "Meaning and Speech Acts". *Philosophical Review* 79, 3—24.
Saver, Jeffrey and Damasio, Antonio (1991). "Preserved Access and Processing of Social Knowledge in a Patient with Acquired Sociopathy Due to Ventromedial Frontal Damage". *Neuropsychologia* 29, 1241–9.
Schachter, S. and Singer, L. F. (1962). "Cognitive, Social and Psychological Determinants of Emotional State". *Psychological Review* 69, 379—399.
Schelling, Thomas (1960). *The Strategy of Conflict* (Cambridge, MA: Harvard University Press).
Trilling, Lionel (1972). *Sincerity and Authenticity* (Oxford: Oxford University Press).

Department of Philosophy
University of Michigan
Ann Arbor, MI 48109
USA