Reason in Philosophy

Animating Ideas

Robert B. Brandom

The Belknap Press of Harvard University Press
Cambridge, Massachusetts · London, England
2009

Contents

Acknowledgments ix

Introduction 1

PART ONE

Animating Ideas of Idealism: A Semantic Sonata in Kant and Hegel

ONE Norms, Selves, and Concepts 27

TWO Autonomy, Community, and Freedom 52

THREE History, Reason, and Reality 78

PART TWO

Reason and Philosophy Today

FOUR Reason, Expression, and the Philosophic Enterprise 111

FIVE Philosophy and the Expressive Freedom of Thought 130

SIX Why Truth Is Not Important in Philosophy 156

SEVEN Three Problems with the Empiricist Conception of Concepts 177

EIGHT How Analytic Philosophy Has Failed Cognitive Science 197

Name Index 227

Subject Index 229

Introduction

This book belongs to a venerable tradition that distinguishes *us* as rational animals, and *philosophy* by its concern to understand, articulate, and explain the notion of *reason* that is thereby cast in that crucial demarcating role. We may call this line of thought "philosophical rationalism." Rationalism of this stripe was not much in favor in Anglophone circles during the last century or so. The predominant tendency in analytic philosophy has been strongly empiricist, at least since Ayer and Carnap. The American pragmatism of James and Dewey defined itself by opposition to a pernicious rationalistic intellectualism. Both those movements, and their heirs in the second half of the twentieth century, were properly recoiling from Cartesian non-naturalism about human minds, knowledge, and agency.

The rationalism that is articulated, motivated, and explored in these pages looks back to Kant and Hegel as its forebears, and to Descartes, Spinoza, and Leibniz only as their deepest lessons came to be understood within that German Idealist tradition. Of course all the figures I invoked in the previous paragraph were reacting against what they understood of that German Idealist tradition, too. (If anything was worse than rationalism, it had to be idealism.) But I think Kant and Hegel showed us a way forward for a rationalism that is not objectionably Cartesian, intellectual-

ist, or anti- (or super-)naturalist. (Nor need it treat the "light of reason" as unacquired or innate.) Recovering the possibility of such a progressive rationalism requires reappropriating central thoughts of Kant and Hegel, disentangling them from the adventitious accretions and subsequent folklore that have grown up so thick around them as almost totally to obscure them. The aim of this book is to make a start on that enterprise, on sketching the sort of rationalism about mind and meaning that emerges from doing so, and on doing that in a way that makes evident its promise and utility for addressing a variety of important philosophical issues that matter to us today.

One of the master ideas pursued here is that rationality is a *normative* concept. In one sense, this is not at all a surprising claim. One can be more or less rational. And other things being equal, more is better. One might object that 'reasonable' and 'rational' are not synonyms in English. Being relentlessly, excessively, or inappropriately rational can be a way of being *un*reasonable. (Just ask anyone who lives with a philosopher!) Perhaps it is better to be more reasonable than to be more rational. But even so, and even though these are different dimensions of normative appraisal, judgments of how rational a belief, commitment, action, or person is do nonetheless have normative consequences. They are relevant to assessments of whether things are as they ought to be, and of what ought or ought not to have been done. (The fact that non-normative descriptions may provide sufficient reasons for distinguishing more from less rational conduct does not speak against the normative significance of such discrimination. Calling something a "good knife" may be wholly warranted by its satisfying such descriptions as "sharp," "durable," and "firmly grippable." It is still a normative characterization.)

This evaluative or comparative normative dimension of rationality rests on a conceptually prior constitutive one. The constitutive issue concerns whether one is a rational creature at all (something that does not come in degrees), rather than whether one is better or worse, more or less reliable, at doing what rational beings as such do. To be a rational being in this sense is to be subject to a distinctive kind of normative appraisal:

assessment of the *reasons* for what one does—in the sense of "doing" that is marked off by its liability to just that sort of appraisal. Rational beings are ones that *ought* to have reasons for what they do, and *ought* to act as they have reason to. They are subjects of rational obligations, prohibitions, and permissions.

It is only creatures that are in the space of reasons in this sense—ones for whom the question of what attitudes they have reason to adopt and what they have reason to do arises, or to whom demands for reasons are appropriately addressed—that are then further assessable as to how *sensitive* they are in fact to their reasons, how good they are at actually doing what they have the best reason to do. Constitutive appraisals of rationality are not *wholly* independent of evaluative appraisals. If some creature *never* acts as it has reason to, is *entirely* unmoved by reasons, is *completely* insensitive to relations of rational consequence and incompatibility among its attitudes, goals, and performances, there might simply be no point in holding it rationally responsible, in treating it as a rational agent and knower at all. Still, there is no definite threshold of heedlessness to reasons that in principle precludes assessment of reasons—of what someone is rationally committed or entitled to.

Taking something to be subject to appraisals of its reasons, holding it rationally responsible, is treating it as some*one:* as one of *us* (rational beings). This normative attitude toward others is *recognition,* in the sense of Hegel's central notion of Anerkennung (discussed in the second chapter of Part One). Adopting that attitude is acknowledging a certain kind of *community* with the one recognized. It is the fellowship of those we acknowledge not only as *sentient* (a factual matter of biology), but also as *sapient* (a normative matter of responsibility and authority). It is attributing a kind of rational personhood, treating others as *selves,* in the sense of knowers and agents, ones who are *responsible* for their doings and attitudes. What they are principally responsible for is having reasons for those doings and attitudes.

The core of the notion of recognition is the idea that normative statuses such as rational responsibility and authority (commitment and entitlement) are fundamentally *social* statuses. They are in principle unintelligible apart from consideration of the practical attitudes of those who *hold* each other responsible, *acknowledge* each other's authority, *attribute*

commitments and entitlements to one another. Because the space of reasons is a *normative* space, it is a *social* space.

Not all norms are rational norms. What is it for it to be *reasons* that one is sometimes obliged or responsible for having, that some sort of authority is conditioned on? The answer developed here understands reasons in terms of *inferences.* Reasons are construed as *premises,* from which one can draw *conclusions.* Although this is, I think, a natural enough approach, it is worth noticing that it is not obligatory. Work is required to wrestle our talk of reasons around into inferential shape. If I am asked why I stacked the flowerpots next to the wall, I may reply "Because she asked me to." That is giving a reason. But it is not immediately clear what relation it bears to any inference. Indeed, when any *fact* is cited as a reason (whether as the justification of a cognitive or a practical commitment), a story would need to be told about how that fact is supposed to be understood to show up as a *premise* in some inference. Inferring is reasoning, but there may be other kinds of reasoning besides inferring. Activities such as making distinctions or comparisons, exploring analogies and disanalogies, conceptual construction and analysis also seem to be kinds of reasoning. It is at least not obvious that they should be either assimilated to or understood in terms of inference. Even in the light of these considerations, I am concerned to see what sort of story can be told, what sort of illumination one can get, by focusing to begin with on the central inferential kind of reasons and the dimension of reasoning they pick out.

One important issue in the vicinity concerns the *semantic* presuppositions of one's account of reasoning and rationality. A methodological fork in the road is marked out by whether one is prepared to take for granted the *meaning* or *content* of whatever serves as or stands in need of reasons. One program that is committed to this order of explanation is rational choice theory (for instance, in its paradigmatic decision-theoretic and game-theoretic forms). It assumes as inputs probabilities and utilities assigned to definite propositionally contentful possible states of affairs. These are part of the specification of the choice situation. It then yields as outputs verdicts about the comparative rationality of various choices and strategies. An explanatory strategy of this shape assumes that the semantic contents of states such as beliefs and desires, contents that can be expressed by declarative sentences, can be settled independently and in ad-

vance of thinking about their role in reasoning. It then addresses questions of the form: what conclusions (practical and theoretical) do commitments of this sort rationally entitle us to? How good are the reasons they give us to do or think one thing rather than another?

Rational choice theory has no indigenous semantics. It outsources that job. Such a division of theoretical labor makes sense, so long as the assumption of independence it presupposes is well-founded. But what if it is not? What if the questions of what a sentence means and what its role in reasoning is are two sides of one coin, needing to be addressed together? Then a different methodological strategy is called for.[1] The broadly Carnapian picture has two stages: semantic and epistemic. First, *meanings* must be fixed: a language. Then, and on that basis, *beliefs* are to be settled: a theory, what sentences with those meanings one has reason to endorse or commit oneself to. This procedure makes sense for artificial languages, but it is arguable that for natural languages there is no way to separate the semantic and the epistemic tasks—not just in practice, but in principle. An alternative Quinean approach does not attempt to address the question of what is a reason for what, independently of the question of what we have reason to believe. The semantic and epistemic dimensions of thought and language use are not only understood as inextricably intertwined, their common structure is the *inferential* articulation characteristic of the space of reasons. In this picture, justification (and so its cousins reason and inference) is not only a key concept in epistemological investigations of the nature of *knowledge,* but also and equally a key concept in semantic investigations of the nature of *meaning.*

Another potentially illuminating broad-brushstroke sketch of a historical path that led to the same constellation of ideas portrays it as the culmination of a three-phase process. The classical American pragmatists were properly impressed by the role beliefs play in the explanation of action. With admirable theoretical ambition, they accordingly tried out the idea that one could not only deduce what people *believe* from what they *do,* but also understand the *concept* of belief (including the *contents* of those

1. I have expressed similar concerns about the parallel semantic assumptions implicit in reliabilist epistemologies. See Chapter Three of *Articulating Reasons* (Cambridge, MA: Harvard University Press, 2000).

beliefs) wholly in terms of the concept of agency. The line of thought that issued in contemporary rational choice theory saw that this could not be right. What people do depends not only on what they *believe,* but also on what they *want.* Intentional actions (they were less concerned than the pragmatists with habitual ones) are to be explained by appeal to beliefs *and* desires, preferences, or some other sort of pro-attitude. (I would prefer to say: "by appeal to *practical* as well as *cognitive* commitments.") The pragmatists show up as having taken for granted a background of needs and wants that played no explicit role in their theorizing. But since beliefs can be held constant and actions varied by varying the accompanying goals,[2] it can be seen that a more articulated scheme is required. Manifesting the same admirable theoretical ambition as the pragmatists, some rational choice theorists likewise proposed turning this model of the explanation of action around and both inferring subjective probabilities and preferences from choice behavior and understanding the notions of probability and preference in terms of their role in explaining choice behavior. The third, Davidsonian phase of this development then consists in pointing out that unless one excludes *verbal* behavior from the scope of one's explanatory ambitions (a drastic, unmotivated, and unnecessary restriction—one that would be particularly bizarre for a theory aimed at *rational* agency), a *third* factor that is being taken for granted is the *meanings* expressed by the sentences the agent utters, including those by which he makes his beliefs and preferences explicit. For once again we could hold fixed an agent's subjective probabilities and preferences and change what she would say and do (for instance, in response to the utterances of others) by varying the meanings she associates with her sentences—most evidently, the sentences expressing the propositions to which she attaches probabilities and preferences. Davidson, too, shares the aspiration of exploiting this form of explanation of action (now tripartite: appealing to beliefs, desires, and meanings) in the service of an order of explanation that starts with specifications of what people *do* and on that basis not only

2. In some contexts, it makes a lot of difference whether one thinks of the practical component in terms of desires, preferences, values, practical commitments, or goals (to mention just a few of the prime contenders). I suppress such complications here in the interests of generality, and because they do not make a difference to the points I am making about various general structures of explanation.

imputes beliefs, desires, and meanings, but also explains theoretically what beliefs, desires, and (especially) meanings *are.*

One lesson we should draw both from Davidson's development of Quine's criticisms of Carnap and from his own criticisms of the theoretical shortcomings that remain, even after rational choice theory substantially improved on classical pragmatist ideas, is that the *inferential* relations sentences stand in to one another are an essential element of the *meanings* that they express. This is a point we can also come at from another direction. Wilfrid Sellars says:

> . . . although describing and explaining (predicting, retrodicting, understanding) are *distinguishable,* they are also, in an important sense, *inseparable.* It is only because the expressions in terms of which we describe objects, even such basic expressions as words for perceptible characteristics of molar objects, locate these objects in a space of implications, that they describe at all, rather than merely label. The descriptive and explanatory resources of language advance hand in hand.[3]

Labels distinguish things. If two objects have different labels, we may conclude that they are different in some respect. But a *mere* label tells us nothing about *which* respect. If two objects have different *descriptions,* however, we not only may conclude *that* they are different, but can consult the *content* of the description to learn something about *how* they are different. Sellars's claim is that we cannot understand that descriptive content apart from the "space of implications" in which the descriptive terminology is embedded. That is, in order to understand the content of the description, we must know something about what other descriptive contents its applicability gives us sufficient reason to apply, which other descriptions would give us reason to apply that one, and which further descriptions it rules out. 'Red' is a description, and not just a label, in part because **being red** follows from **being scarlet,** entails **being colored,** and

3. "Counterfactuals, Dispositions, and the Causal Modalities," in *Minnesota Studies in the Philosophy of Science,* vol. 2, ed. H. Feigl, M. Scriven, and G. Maxwell (Minneapolis: University of Minnesota Press, 1957), sec. 108.

rules out **being green.** Understanding a description, as opposed to being able to apply a label, is a matter of practically mastering the *inferential* relations it stands in to other descriptions: its place in the space of reasons or implications. The parrot we have trained reliably to respond to the visible presence of red things by squawking "That's red" is applying a label. A three-year-old child who knows that red lollipops have a cherry taste and red traffic lights mean "Stop" is already applying a descriptive *concept.*

So one of the things I take it we have learned is that *reasoning* and *concept use* are two sides of one coin. *Discursive* activity, applying concepts paradigmatically in describing how things are, is inseparable from the *inferential* activity of giving and asking for *reasons.* Elsewhere[4] I have elaborated a view along these lines under the rubric "semantic inferentialism." Three ever more committive versions of such a view can be distinguished. *Weak* semantic inferentialism claims only that inferential articulation is a *necessary* condition of conceptual contentfulness. *Strong* semantic inferentialism claims further that inferential articulation in a broad sense is a *sufficient* condition. *Hyper*inferentialism claims that inferential articulation in a narrow sense is a sufficient condition. In *Making It Explicit* I developed and defended a version of strong semantic inferentialism. In this book, I explore some of the motivations and consequences of such a view.

Understanding conceptual content in terms of role in reasoning, and reasoning in terms of inference, entails giving pride of place to *propositional* content. For propositional contents, in the first instance expressed by declarative sentences, can play the role both of premise and of conclusion in inferences—and so, on the inferential conception of reasoning, both serve as and stand in need of reasons. When I was in elementary school, I was solemnly taught that "a sentence is the grammatical expression of a whole thought." Even at the time, I found the confident endorsement of this order of explanation quaint and unmotivated. Surely the two ideas *sentence* and *whole thought* come as a package. The issue being implicitly raised is the vexed philosophical topic of the "unity of the propo-

4. *Making It Explicit* (Cambridge, MA: Harvard University Press, 1994) and *Articulating Reasons*.

sition." What is the unity characteristic of basic thinkables (namely judgeables), as such? The suggestion is that this question, too, should be addressed in terms of role in *reasoning*—and hence, given the specifically inferentialist version of rationalism being pursued here, *inference.* What is propositionally contentful is what can stand in inferential relations. (In this sense, it is not only *declarative* sentences that exhibit propositional contentfulness. Expressions of intention or preference, requests, and commands can do so as well.)

What is the relation between reasoning (inferring) and inferential relations (among propositionally contentful items)? A traditional strategy seeks to derive proprieties of inferential practice from inferential relations. But there are reasons to adopt a converse strategy. Gilbert Harman has famously and provocatively argued that there is no such thing as deductive inference.[5] If there were, he suggests, then it would be conducted according to rules such as "If you believe that *p,* and you believe that *if p then q,* then you should believe that *q.*" But that would be a terrible rule. One might have much better evidence against *q* than one has in favor of either *p* or *if p then q.* Deductive implicational relations tell us that we ought not to believe all of *p, if p then q,* and *∼q.* But they don't tell us what to *do,* which of the many options for repairing the situation to pursue, if we are threatened by commitment to such an inconsistent triad. Inferential relations constrain, but do not determine, inferential processes.

A *pragmatist* strategy looks to impute inferential *relations* on the basis of inferential *processes* or *practices.* (Think here of the heroic functionalist inversion strategy mentioned earlier: the idea of turning on its head the form of an explanation of action in terms of reasons—whether in terms of beliefs, beliefs plus desires, or beliefs plus desires plus meanings—so as to explain the nature of the explainers, including their contents, in terms of their role in such explanations.) The fundamental termini of inferring moves—the acts or statuses that are givings of reasons and for which reasons are given—are judgings, claimings, assertings, or believings. They are the undertakings or acknowledgments of *commitments.* It is the propositional contents of those commitments that stand in inferential relations

5. "Logic and Reasoning," *Synthese* 60 (1984): 107–128.

to one another. The indissoluble package that includes sentence on the syntactic side, and propositional content on the semantic side, also includes the activities of inferring and (therefore) asserting on the pragmatic side. This is the iron triangle of discursiveness.

On this conception of the discursive, then, to say that we are discursive creatures is to say both that we are *normative* creatures and that we are *rational* ones. We undertake discursive commitments and responsibilities, and what makes them *discursive* commitments and responsibilities is that they stand in *inferential* relations to one another: relations that codify what is a *reason* for what. Because our commitments are inferentially articulated, they are *conceptually* contentful. The space of reasons is the space of concepts. What discursive beings as such do is apply concepts, undertaking cognitive commitments as to how things are and practical commitments as to how things shall be. Such discursive activity is the exercise of a distinctive kind of *consciousness.* It is *sapient,* rather than merely *sentient,* consciousness or awareness. For it depends on the sort of conceptual *understanding* that consists in practically knowing one's way about in the inferentially articulated space of reasons and concepts, rather than the sort of organic *feeling* we share with animals that are not *rational* animals.

This rationalist understanding of our characteristic discursive kind of consciousness also makes sense of a corresponding sort of semantic *self*-consciousness. For we can take things to be thus-and-so, *describe* them (rather than merely discriminate or label them), become discursively aware of them, in virtue of the inferential relations our commitments (including those we come to acknowledge by non-inferential responsive perceptual mechanisms) stand in to one another. And those same inferential relations can themselves become the topics of our discursive awareness. All that is required for that is practical mastery of the use of vocabulary expressing the right kind of concepts. These are concepts that let us make explicit—put into judgeable, thinkable, assertable, *propositional* form—the inferential relations that articulate the "space of implications" that is the context and horizon within which alone what we do acquires the significance of rational, discursive consciousness of what we respond to. These are concepts that let us *say that* one claimable content follows from or is a reason for or against another. The paradigmatic locution express-

ing such a concept is the *conditional.* "If *p* then *q,*" tells us that a certain inferential relation holds between *p* and *q.*

In the rationalist picture I have been sketching, inferential relations are the medium of discursive understanding and consciousness in the sense of sapience. The expressive role characteristic of *logical* vocabulary is to codify inferential relations in propositional form. That is to bring them into the practices of giving and asking for reasons as themselves items that can be given as reasons and for which reasons can be sought—that is, as things that can themselves be objects of discursive understanding and consciousness. Where before the advent of logical vocabulary and the concepts it expresses rational creatures had to be able in practice to discriminate (however fallibly and incompletely) materially good from materially bad inferences, the new expressive capacity provided by logical locutions lets them reason about what is a reason for what. Logic is the organ of semantic self-consciousness.[6]

So one consequence of this way of thinking about things (a potentially controversial one, to be sure) is that one need not be a *logical* being to be a *rational* one. The sort of semantic (because inferential) *self*-consciousness that is afforded by mastery of the use of specifically logical vocabulary is not, as far as I can see, a necessary condition of the sort of sapient consciousness afforded by mastery of the use of ordinary empirical descriptive vocabulary. One can be good enough at distinguishing what follows from what and what is a reason for or against what other claims to be held responsible for having reasons for one's cognitive and practical commitments even if that ability remains at the implicit level of knowing *how,* and is not something one could formulate explicitly in rules or principles at the level of knowing (and being able to say) *that* certain inferences are good and others not. The sort of self-consciousness that is exhibited in making *explicit* (in judgeable, sayable, thinkable, *propositional* form) what otherwise remains *implicit* in the inferentially articulated practical capacities in virtue of which one can be consciously aware of anything (make something explicit, by judging, saying, or thinking

6. This is the doctrine that I call "logical expressivism" in *Making It Explicit,* chap. 2, and *Articulating Reasons,* chap. 1, and have articulated in most detail in *Between Saying and Doing* (Oxford: Oxford University Press, 2008), chaps. 2 and 5.

something, applying concepts) is *sui generis,* and not to be understood on the traditional Tarski-Carnap model of metalanguages. Logical vocabulary and concepts make possible a distinctive kind of self-reflection about reasons and reasoning, and so about the semantogenic features characteristic of specifically discursive practice.

This is to say that philosophy begins in logic. For as I remarked in the opening sentence of this Introduction, in the broadly rationalist tradition to which this work belongs, philosophy is demarcated by its concern to understand, articulate, and explain the notion of *reason* that distinguishes us as *rational* animals, *discursive, concept-using, sapient* beings. Specifically *logical* self-consciousness is a matter of being able to make claims and reason about reasoning, about inference and the inferential relations that articulate the contents of *non*-logical concepts. So logic makes possible already a kind of distinctively *philosophical* reflection.

Of course there is more to philosophy than logic. Put slightly more carefully, not all the concepts whose use is the exercise of specifically philosophical self-consciousness (consciousness of ourselves as *rational* animals) are logical concepts. I have emphasized that on my way of understanding it, rationality is a normative concept. The space of reasons is a normative space. I have also registered a methodological commitment (though in these introductory remarks I have not tried to entitle myself to it by argument) to a pragmatist order of explanation—to accounting for *meaning* in terms of *use;* more specifically, to abstracting the inferential relations that articulate conceptual (paradigmatically propositional) content from the reasoning practices and inferential processes of discursive practitioners. This is a broadly functionalist explanatory strategy. It is a rationalist (more specifically, inferentialist) functionalism, because conceptual content is to be understood in terms of role in reasoning, in the form of inferential role. And it is a normative functionalism, rather than a causal-dispositional functionalism. That is, the roles in question are to be specified in a normative vocabulary of what would commit or entitle one to apply a concept and what else doing that would commit or entitle one to, rather than with what would dispose one to apply that concept and what else doing that would dispose one to do. (For reasons discussed in Chapter Two, my understanding of normativity then requires a social

functionalism, in which the functional system in question is a linguistic community and its practices, rather than confining the functional system to what is inside the skin of some individual organism. But that further commitment can be ignored here.)

So according to an account that founds an inferentialist semantics on a normative pragmatics, philosophical reflection on us as rational creatures must deploy not only *logical,* but also *normative* concepts. The latter have the expressive job of making explicit what it is we are *doing* in engaging in discursive practices, applying concepts, and exercising our sapient consciousness. For what we are doing is claiming authority and undertaking responsibility, altering our commitments and entitlements in ways that depend on what is a reason for what. Besides developing and deploying logical concepts and vocabulary to express inferences in explicit, propositional, conceptual form, philosophers must develop and deploy normative concepts and vocabulary to express the acknowledgment and attribution of discursive deontic statuses such as responsibility and authority, commitment and entitlement. Philosophy is concerned not only with the semantics, but with the pragmatics of discursivity. That much follows already from the most basic characterization of its task as understanding us as rational creatures, once the normativity of that characterization is appreciated.

This constellation of ideas provides the framework within which the more specific investigations reported in the body of this work are conducted. The book has two parts. The first finds in normative rationalism an organizing principle that guides us to a new way of understanding both Kant's most basic ideas and what Hegel made of those ideas. These three chapters are tightly integrated with one another, presenting a single, cumulative story culminating in a unified, articulated picture of the rational activity that makes us what we are: knowers and agents, creatures able to undertake and attribute determinate, conceptually contentful commitments and responsibilities. The second part of the book comprises five chapters that discuss various topics in contemporary philosophy, viewed from the perspective of what it has to tell us about ourselves as rational creatures.

Part One of the book consists of my three Woodbridge lectures.[7] The first of these, "Norms, Selves, and Concepts," begins with what I take to be Kant's principal innovation: his normative characterization of the mental. This is the idea that what distinguishes judging and intentional doing from the activities of non-sapient creatures is not that they involve some special sort of mental processes, but that they are things knowers and agents are in a distinctive way *responsible* for. Judgments and actions make knowers and agents liable to characteristic kinds of *normative* assessment. What one must *do* in order to be taking responsibility for or committing oneself to a judgeable content (or practical maxim) in the sense that matters for apperceptive (sapient) awareness is *synthesize* an original unity of apperception, by *integrating* the content in question into the whole that comprises all of one's commitments, in the light of the relations of material inferential consequence and incompatibility they stand in to one another. This is the synthesis of a rational self or subject: what is responsible *for* the commitments. It has a *rational* unity in that the commitments it comprises are treated as *reasons* for and against other commitments, as normatively *obliging* one to acknowledge some further commitments and *prohibiting* acknowledgment of others. This is Kant's normative inferential conception of awareness or experience. Further, he pursues a pragmatist order of explanation, in that concepts and their contents are understood in terms of the role they play in this synthetic process of taking rational responsibility (judging). As predicates of judgment, concepts are rules determining what is a reason for what. The final move in this first chapter is to explain in these terms the notion of representation. Kant understands this, too, in normative terms. What is represented is what representings (judgments) are responsible *to,* what exercises a distinctive kind of *authority* over assessments of their correctness. I offer an account of this sort of normativity in terms of the more basic inferential sort that governs the rational synthetic process, and so provide a reading

7. Originally presented at Columbia University in November of 2007. I delivered them again at the University of Pittsburgh, in January of 2008, and once more at the Humboldt University in Berlin, in July of 2008. I have profited greatly from the comments of all those audiences.

of Kant's dark but central claim that "it is the unity of consciousness that alone constitutes the relation of representations to an object."[8]

The first chapter, then, presents Kant's pragmatist version of a normative rationalist account of intentionality, along both its expressive dimension and its representational dimension. For both propositional (and, more generally, conceptual) contentfulness and what it is for our judgments to purport to be about objects (which play the normative role of what those judgments answer or are responsible to for their correctness) are explained in terms of the activity of judging, understood as a process of synthesizing a unity of apperception by rational integration of commitments. The second chapter, "Autonomy, Community, and Freedom," explores Kant's theory of *normativity*. (The shift in topic corresponds to that from concern with *rational* normativity to concern with rational *normativity*.) Earlier Enlightenment thinkers (for instance, in the social contract tradition of political thought, culminating in Rousseau) had already had the idea that normative *statuses* such as responsibility and authority (commitment and entitlement) are not independent of the normative *attitudes* of those who acknowledge such responsibility or authority. Kant radicalizes that idea into a criterion of demarcation of the normative—a way of distinguishing normative constraint from various sorts of nonnormative compulsion—in terms of *autonomy*. One is genuinely *normatively* bound only by rules one has bound *oneself* by, concepts one has oneself applied in judging or acting. The central role accorded to normativity in understanding concept use generally (in Chapter One) turns out here to have as a consequence that the significance of the concept of *autonomy* for Kant's thought extends far beyond the practical use of reason, penetrating deeply into his account of rationality and discursivity in general—in their theoretical no less than their practical forms of expression.

At this point Hegel's reciprocal recognition model of normativity is introduced. It takes its place as a way of developing Kant's autonomy model, so as to make sufficient room for the relative *independence* of conceptual contents from the attitudes of those who endorse them. That independence is in turn required to make intelligible the notion that in applying concepts one is undertaking a determinately contentful rational responsi-

8. *Critique of Pure Reason* B137.

bility. Hegel's social model manages to satisfy that criterion of adequacy while still respecting the original Enlightenment insight. It does so by acknowledging the crucial role played by the attitudes of *others*—those I recognize (as authoritative) and am in turn recognized by—in constituting the authority I have to make myself responsible to conceptual norms.

Kant introduced a new normative conception of positive freedom. Freedom in his sense consists in the capacity to bind oneself by norms. It is the authority to make oneself responsible. The chapter closes with a discussion of how this looks when it is transposed into Hegel's new social key, and given, as it is by Hegel, a further, specifically *linguistic* twist. The positive expressive power achievable only by constraining oneself by the norms constitutive of the use of a public natural language provides a paradigm of the Hegelian version of Kant's normative conception of positive freedom.

The third chapter, "History, Reason, and Reality," describes how Hegel fits together the model of the synthesis of an original unity of apperception by rational integration (discussed in the first chapter) and the model of the synthesis of normative-status-bearing apperceiving selves and their communities by reciprocal recognition (discussed in the second chapter) by placing both within a larger *historical* developmental structure. Further, he does so in such a way as to make intelligible the sense in which the discursive commitments that arise in that process should be understood as having determinate conceptual contents. Conceptual contents are determined by rational integration that includes historical reflection. The Kantian account of rational integration of new commitments into a synthetic unity with prior commitments is recontextualized as merely one aspect of a more general rational integrative-synthetic activity. What Hegel adds is a *retrospective* notion of *rationally reconstructing* the process that led to the commitments currently being integrated (not just the new one, but all the prior ones that are taken as precedential for it, too). This is a kind of *genealogical* justification or vindication of those commitments, showing why previous judgments were correct in the light of still earlier ones—and in a different sense, also in the light of subsequent ones. Hegel calls this process "Erinnerung," or *recollection.* A successful recollective reconstruction of the tradition shows how previously endorsed constellations of commitments were unmasked, by internal in-

stabilities, *as appearances,* representing how things *really* are only incompletely and partially incorrectly—but also how each such discovery contributed to filling in or correcting the picture they present of how it really is with what they were all along representing, by more closely approximating the actual consequential and incompatibility relations of the concepts and making more correct applications of them. So they were not *mere* appearances, in that they did genuinely reveal something of how things really are. Exhibiting a sequence of precedential concept applications-by-integration as *expressively progressive*—as the gradual, cumulative making explicit of reality as revealed by one's current commitments, recollectively made visible as having all along been implicit—shows the prior, defective commitments endorsed, and conceptual contents deployed, as nonetheless genuinely appearances representing, however inadequately, how things really are.

The new notion of *reason,* expanded to include both *integration* and *recollection,* is the centerpiece of an account of what discursive practitioners must *do* in order to be intelligible as granting *authority* over the correctness of what they say and think (in a sense of 'correct' corresponding to a distinctive normative dimension of assessment they institute by those very practical attitudes) to an objective reality they count, by making themselves *responsible* to it, thereby in this normative sense as talking and thinking *about.* In fact the systematic account of reason and concepts, normativity and freedom, of the social and historical dimensions of self-consciousness, and of intentionality and objectivity that is recounted in these three chapters itself exhibits this same structure. These chapters recollectively reconstruct an expressively progressive rational path leading from Kant's central ideas to Hegel's. They exhibit Kant and Hegel as developing a sophisticated account of the unity of the various central features that distinguish discursive beings from merely natural ones. We are social, normative, rational, free, self-consciously historical animals. This powerful and finely conceptually articulated vision of ourselves is the crowning philosophical achievement of German Idealism.

The material in Part Two of the book is less intricate and more accessible. These chapters should be intelligible not only to philosophers, but also to more general audiences. They try to extract lessons that might matter in other corners of the culture from the particular way of working

out the rationalist tradition of thinking about us that I have been waving my hands at. They take this same constellation of considerations developed in Part One and apply them to more specific issues of contemporary philosophical significance.

Chapter Four, "Reason, Expression, and the Philosophic Enterprise," begins where Chapter Three left off: considering us as creatures with histories rather than natures. It then takes a metaphilosophical turn, addressing the question of what philosophy is. The overall view is that it is the discipline that thinks about what *we* are, first as beings who live and move and have our being in a *normative* space, and second as *rational, discursive* creatures. Its conclusion is that the topic of philosophy is normativity in all its guises, and inference in all its forms. And its task is an *expressive, explicative* one. Philosophy is a discipline whose distinctive concern is with a certain kind of *self-consciousness:* awareness of ourselves as specifically *discursive* (that is, concept-mongering) creatures. Its task is understanding the conditions, nature, and consequences of conceptual norms and the activities—starting with the social practices of giving and asking for reasons—that they make possible and that make them possible. As concept users, we are beings who can make explicit how things are and what we are doing—even if always only in relief against a background of implicit circumstances, conditions, skills, and practices. Among the things on which we can bring our explicitating capacities to bear are those very concept-using capacities that make it possible to make anything at all explicit. Doing that is philosophizing.

Chapter Five, "Philosophy and the Expressive Freedom of Thought," takes a different path through what should now be familiar material. It, too, is metaphilosophical in intent.[9] The argument proceeds by way of a consideration of the relative roles of our sentience and sapience in thinking about who *we* are—in our ownmost being, as Heidegger liked to say. Three models of the sapience of rational animals are considered: an instrumental one, which considers rationality as a matter of being good at

9. It was originally written as a contribution to a popular lecture series at Yale in which different advocates spoke in favor, respectively, of a life of pleasure, of philosophical activity, and of philosophical reflection. Their original agreed-upon avatar for the life of pleasure was Mick Jagger—but at the last minute he begged off. Something about a banquet, I understand.

getting what we want, a Kantian normative one, and a Hegelian social-historical one. The conclusion, for what it is worth, is the Hegelian one: that what we are *in* ourselves depends on what we are *for* ourselves. We are the kind of being whose self-conception is essential to our selves. These different self-conceptions correspond to different aims for political activity. And they correspond to different conceptions of philosophy. The piece ends with an encomium to the philosophical life. It is not meant, *pace* Rousseau, to claim that one must first be a philosopher in order truly to be one of us,[10] only to recommend the special virtues of the special form of self-consciousness that philosophers (those who just can't help doing it) aim to practice and promote.

Chapter Six, "Why Truth Is Not Important in Philosophy," addresses a question that might well have occurred to the reader very early on in the exposition of this Introduction: why reason and not truth as the master idea in terms of which to approach both the question of who we are and the question of what philosophy is and by rights ought to be? After all, the two concepts have equally august pedigrees in the philosophical tradition. As the point is put at the end of this chapter, we are not only "makers and takers of reasons," but "seekers and speakers of truth." The methodological route I took here looks first to our sapience, and then to a specifically *semantic* account of the conceptual contents it is our fate, mission, and glory as sapients to shape and be shaped by. But that very sort of semantic story is, in the recent tradition, typically told in terms of truth and truth conditions, rather than reason and inference (which come into the story only later). I am concerned here (as in *Tales of the Mighty Dead*) retrospectively to discern and provide a provenance for a contrary explanatory strategy. But surely its competitor should at least be addressed.

My general view is that once the *expressive* role distinctive of the vocabulary of 'truth' and 'reference' is properly appreciated, it will be seen that the concepts expressed by these locutions are not suitable to be recruited for semantic *explanatory* purposes. This is a conclusion that my allies in the semantic deflationist camp have not always sufficiently emphasized. (In the context of the general direction of explanation I have been recommending, such a view will have substantial consequences for

10. I have in mind his remark in the preface to the *Discourse on the Origin of Inequality*.

how one thinks about us as rational animals, sapients, and discursive practitioners—and hence how one thinks about philosophy as the study of us as such.) More deeply, from my pragmatist perspective, I am impressed that inferring and asserting are things that one *does.* (*Representing,* for instance, is only something one does in a seriously derivative, parasitic sense—something one does *by* doing something more fundamental.) Of course, speaking *truly* is also something one does. But 'truly' is an adverb, not a verb. And saying something *true* is also something one does. Yet 'true' is an adjective, not a verb. Obviously, grammar is not destiny. But pragmatists (or just functionalists) want to know how these assessments are related to what people *do* in engaging in discursive practices: in thinking and saying things.

This chapter is divided into two parts. The critical bit describes the expressive role characteristic of semantic vocabulary. Truth-talk generally ought to be understood anaphorically, as a way of forming prosentences, which relate to their antecedents in ways analogous to the ways pronouns relate to theirs. If that is right, then there are various explanatory roles that philosophers have hoped to recruit the concept of *truth* to pursue that it turns out not to be suitable for. The constructive bit of the chapter explores how the notion of *inference* can be used to pursue many of those same philosophical projects for which *truth* traditionally seemed to be essential, or at least helpful.

One cause for complaint about this chapter is that in taking success semantics as a target along the way, I have chosen a minority view that is well out of the mainstream of truth-conditional semantics. That is true. I have seized on it because it seems to me to be one of the most sophisticated truth-conditional approaches that takes broadly pragmatist considerations seriously in its larger philosophical motivations. Overall, my aim is to show how things look from a point of view that emphasizes reason. And I have nothing against working out a different approach. But it is also important to register some of the drawbacks of a popular competing approach.

Chapter Seven, "Three Problems with the Empiricist Conception of Concepts," begins by distinguishing four ways in which a broadly functionalist approach to the conferral of conceptual content might be pursued: the empiricist strategy, which looks exclusively to the causal ante-

cedents that reliably elicit a (therefore) contentful state; the pragmatist strategy, which looks exclusively to the causal consequences reliably elicited by a (therefore) contentful state; the rationalist strategy, which looks to the role in reasoning of a (therefore) contentful state; and the combined strategy, which appeals to all these sorts of functional involvements, both causal and inferential. The aim is to show some of the ways in which the rationalist criterion of demarcation of the *conceptual* in terms of specifically inferential articulation remedies specific shortcomings of empiricist and pragmatist accounts of conceptual content that appeal only to causal inputs (in perception) and outputs (in action). To that end, the chapter considers three difficulties that confront such non-inferential conceptions of conceptual content: inability to distinguish the sort of complexity distinctive of the *conceptual* as such, failure to make sense of the crucial notion of responsive *reliability,* and liability to *proximal* interpretations of what empirical concepts are applied to. In each case, it is claimed, the rationalist insight supplies what is missing from the empiricist approach.

The final chapter, "How Analytic Philosophy Has Failed Cognitive Science," argues that concept use is intrinsically stratified. It exhibits at least four basic layers, with each capacity to deploy concepts in a more sophisticated sense of 'concept' presupposing the capacities to use concepts in all of the more primitive senses. The three lessons that generate the structural hierarchy oblige us to distinguish between

- concepts that only *label* and concepts that *describe,*
- the *content* of concepts and the *force* of applying them, and
- concepts expressible already by *simple* predicates and concepts expressible only by *complex* predicates.

Artificial intelligence (AI) researchers and cognitive, developmental, and animal psychologists need to take account of the different grades of conceptual content made visible by these distinctions.

Considerations that have been in play since the dawn of analytic philosophy, well over a century ago, yield a four-stage hierarchy of ever more demanding senses of "concept" and "concept use." At the bottom are concepts as reliably differentially applied, possibly learned, *labels* or clas-

sifications. This is the empiricist construal of concepts which I argued in the previous chapter needs to be supplemented in order to be adequate. At the next level, concepts as *descriptions* emerge when merely classifying concepts come to stand in *inferential, evidential, justificatory* relations to one another—when the propriety of one sort of classification has the practical significance of making others appropriate or inappropriate, in the sense of serving as *reasons* for them. This is the rationalist construal of concepts.

Building on the capacity to use inferentially articulated descriptive concepts to make propositionally contentful judgments or claims, the capacity to form sentential *compounds*—paradigmatically *conditionals,* which make endorsements of material inferences relating descriptive concept applications propositionally explicit, and *negations,* which make endorsements of material incompatibilities relating descriptive concept applications propositionally explicit—brings with it the capacity to deploy a further, more sophisticated kind of conceptual content: *ingredient* (as opposed to freestanding) content. Conceptual content of this sort is to be understood in terms of the contribution it makes to the content of compound judgments in which it occurs, and only thereby, indirectly, to the force or pragmatic significance of endorsing that content. Ingredient conceptual content is what can be *negated,* or *conditionalized.*

The first step was from merely *discriminating* classification to *rational* classification ('rational' because inferentially articulated, according to which classifications provide reasons for others). The second step is to *synthetic logical* concept formation, in which concepts are formed by logical compounding operators, paradigmatically conditionals and negation. The final step is to *analytical* concept formation, in which the sentential compounds formed at the third stage are *decomposed* by noting invariants under substitution. Systematically assimilating sentences into various equivalence classes accordingly as they can be regarded as substitutional variants of one another is a distinctive kind of *analysis* of those compound sentences, as involving the application of concepts that were not *components* out of which they were originally constructed. Concepts formed by this sort of analysis are substantially and in principle more expressively powerful than those available at earlier stages in the hierarchy of conceptual complexity.

This hierarchy is not a *psychological* one, but a *logical* and *semantic* one. Concepts at the higher levels of complexity presuppose those at lower levels not because creatures of a certain kind cannot in practice, as a matter of fact, deploy the more complex kinds unless they can deploy the simpler ones, but because in principle it is impossible to do so. Nothing could count as grasping or deploying the kinds of concepts that populate the upper reaches of the hierarchy without also grasping or deploying those drawn from its lower levels. The dependencies involved are not empirical, but (meta)conceptual.

The five chapters that make up the second part of this book thus represent a number of different paths that can be taken through what are essentially the same rationalist woods explored in the first three. The desire to make it possible to read them also as self-contained essays means that some prominent features of that landscape show up many times, from somewhat different perspectives. Those landmarks provide orientation, and a potentially calming reminder on unfamiliar ground that at least one already knows what is to be found in a neighboring region. The hope is that by the end of these journeys, the reader will feel thoroughly at home in this terrain.

The book as a whole describes a narrative arc that takes us from Kant to the frontiers of contemporary analytic philosophy. It presents a rationally reconstructed tradition that ties these endpoints together in an expressive progression, in which a promising constellation of ideas gradually becomes more explicit. One of Hegel's big ideas is that this recollective form of rationality is the way one moves forward, by looking backward. (Traditions are lived forward and understood backward.) An understanding of how we got to where we are is a distinctive form of an idea about where we can go. It is of the essence of this form of understanding that it is not unique. There are many different diagnoses possible, leading to different proposed therapies. There are eight million stories out there, and this is one of them.