

## Chapter 3

# How Moral Psychology Changes Moral Theory

*Mark L. Johnson*

---

### *The Moral Philosophy versus Moral Psychology Split*

A great many philosophers think that moral philosophy does not have to pay much attention to moral psychology. They think either that moral psychology is mostly irrelevant to moral theory, or they believe that rational self-reflection alone can generate an adequate set of psychological assumptions without relying on any empirical studies from moral psychology. Moral purists of this sort labor under the illusion that there exists a large gulf separating moral theory from moral psychology. They regard "pure" moral philosophy as being concerned only with how we ought to reason and act and with justifying the fundamental principles of morality. They then contrast this sharply with moral psychology, which they allege to be a merely empirical discipline describing the contingent facts about how people actually are motivated, how they understand things, and the factors that affect their moral reasoning. Armed with this grand distinction between moral theory and moral psychology, along with its attendant assumptions of an is-ought split and a fact-value dichotomy, defenders of a very narrow conception of moral philosophy pretend to dismiss the complex, messy concerns of moral psychology, which are regarded as being irrelevant, or at least tangential, to the tasks of moral theory.

Those who want to deny the relevance of moral psychology to moral theory typically try to make their case by assuming an extremely narrow and trivial conception of moral psychology as being concerned only with the psychological conditions that affect concrete deliberations and decisions within specific situations. Knowing why this or that individual or group reasoned and acted in a certain way, for instance, certainly does not tell whether they acted in a morally praiseworthy manner. Knowing why so many people were attracted to the values and social institutions of nazism does not indeed tell whether those values and institutions were good or bad. Consequently, this trivialized conception of moral psychology can make it seem as though psychology has no important relation to moral theory.

Moral psychology, however, is not psychology in this narrow sense. Rather, moral psychology should be understood broadly as what I will call the psychology of human moral understanding, which includes empirical inquiry into the conceptual systems that underlie moral reasoning. The psychology of moral understanding can give us profound insights into the origin, nature, and structure of our basic moral concepts and into the ways we reason with those concepts. There is no direct deductive link between such knowledge of our moral concepts and specific moral rules (such as rules telling us why nazism is immoral), and that is why moral psychology is not going to give an exhaustive set of prescriptions for moral living. However, moral psychology will tell what is involved in making moral judgments, and it will thereby cultivate in us a certain wisdom that comes from knowing about the nature and limits of human understanding—a wisdom that will help us live morally insightful and sensitive lives.

As a graduate student in the early 1970s, I was indoctrinated, as many other generations of philosophy students have been, with both this trivializing view of moral psychology and also with an extremely restrictive conception of the nature of moral philosophy that has its roots partly in Enlightenment epistemology and partly in the influential pronouncements of G. E. Moore about the nature of ethics. According to the received view, there were alleged to be three radically distinct enterprises that jointly made up the field of ethics:

1. *Descriptive ethics*. This was thought to be a mere empirical investigation of moral standards and practices across times and cultures. As merely descriptive, it was alleged to have no “prescriptive” force, that is, no direct bearing on philosophic attempts to determine how we ought to act.
2. *Normative ethics*. This was supposed to be an attempt to lay down prescriptive moral principles meant to guide our action, our willing, and our moral evaluation of actions and persons.
3. *Metaethics*. This was conceived as a form of conceptual analysis of the cognitive status and semantic content of various moral concepts. Moore’s *Principia Ethica* went a long way toward defining moral philosophy as primarily concerned with clarification and analysis of our fundamental moral concepts, such as good, right, duty, and rule. In itself, metaethics was not believed to be a prescriptive or normative activity, although there were intimations that good conceptual analysis would clarify a number of issues relevant to normative ethics.

The nearly exclusive focus on metaethics that characterized at least the first six decades of this century in Anglo-American philosophy was surely the nadir of moral theory. This impoverished state of moral theory ex-

plains why so many of us felt considerable excitement and liberation in the late 1960s and early 1970s when we first encountered the work of John Rawls. Rawls sidestepped the dominant metaethical questions and went straight to the genuine normative concerns of moral and political theory. His sophisticated nonfoundationalist epistemology seemed to make it possible for us to do normative ethics once again. However, in spite of this new wave of constructive moral theory, the split between moral philosophy and moral psychology was never seriously questioned. Rawls said only that any adequate moral theory must be generally compatible with our most reliable theories of moral psychology. He did not, however, give a central role to questions of moral psychology, and this has left the moral theory versus moral psychology dichotomy relatively intact for much of the last quarter-century.

The obvious question that arises is why this split between moral theory and moral psychology runs so deep in twentieth-century moral philosophy. The answer, I believe, is tied up inextricably with our traditional conception of the role that moral philosophy ought to play in our lives. People tend overwhelmingly to want their moral philosophy to give them moral guidance for their lives. People want a rational way—a method—for determining how they ought to act in the kinds of situations they typically encounter. A moral theory that does not lead fairly directly to prescriptions for action would be, on this account, no moral theory at all.

This desire for a moral guidance is natural and understandable enough, given the complexity and indeterminacy of our human experience. However, it is a short misstep from this reasonable desire for moral guidance to the claim that any satisfactory moral theory ought to be what I call a "governance" theory, that is, one that gives a set of moral rules specifying how we should act in concrete situations. In contemporary theory, Alan Donagan has articulated a prototypical governance conception, in which morality is defined as "a standard by which systems of mores, actual and possible, were to be judged and by which everybody ought to live, no matter what the mores of his neighbors might be." Moral theory, on this view, is thus "a theory of a system of laws or precepts, binding upon rational creatures as such, the content of which is ascertainable by human reason."<sup>1</sup>

Donagan's theory is quintessentially a governance theory, one that purports to set out definite moral rules that specify how we ought to deal with the complex moral problems of contemporary life. Like Rawls and most other important contemporary moral theorists, Donagan would certainly have thought that no moral theory can be acceptable that presupposes a view of human psychology that is demonstrably false. Almost everyone today would say, at least, that we cannot ignore moral psychology in our moral philosophizing.

Still, many moral philosophers do ignore what is going on in moral psychology, and in the cognitive sciences generally. It is a depressing fact that the moral theory versus moral psychology split still stands, for theoretical reasons or simply because psychology continues to be neglected by philosophers. The impoverished state of twentieth-century moral philosophy is shown by the fact that we so desperately needed a book like Owen Flanagan's *Varieties of Moral Personality*, which devotes 400 pages to defending the proposition that moral theory must incorporate a realistic human psychology. Flanagan has taken the first major steps in setting out the general parameters of an adequate moral theory that would satisfy what he calls the Principle of Minimal Psychological Realism: "Make sure when constructing a moral theory or projecting a moral ideal that the character, decision processing, and behavior prescribed are possible, or are perceived to be possible, for creatures like us."<sup>2</sup> In other words, only recently Flanagan found it necessary to mount a major offensive in order to make the world safe for moral psychology. It is a good and important book, but he should not have been forced to spend so much energy reminding us of the necessity of a psychologically realistic moral theory.

Even more recently, Samuel Scheffler has emphasized "the importance for moral philosophy of some tolerably realistic understanding of human motivational psychology," explaining that he is "convinced that the discussion of some of the central questions of moral philosophy could only benefit from a more serious attention to psychological reality."<sup>3</sup> Unfortunately, Scheffler proceeds to ignore, for the most part, the large body of work in moral psychology that could be relevant to his argument. Feminist moral philosophers have also been arguing for a psychologically realistic moral theory for well over a decade, uncovering a network of assumptions that underlie the dominant conception of moral theory, with all its foundational dichotomies and gendered concepts. And we can, of course, go back to philosophers like James and Dewey for robust moral psychologies that could provide the basis for realistic moral theories.

What I have been urging so far is this: many philosophers cling tenaciously to the moral philosophy versus moral psychology split because they think that this is the only way to preserve a governance theory of morality and, with it, the idea that moral philosophy can give us moral guidance. They fear that wading into moral psychology can only teach us how and why people do what they do, but without telling us definitively what people ought to be doing.

I am going to argue that the central purpose of moral theory should be the enrichment and cultivation of moral understanding and that moral psychology is essential to the development of our moral understanding. Moral psychology therefore lies at the heart of any adequate moral theory. There will be moral guidance from moral theory so construed, but only of

the sort that comes from moral insight into complex situations and personalities, rather than the sort that comes from applications of moral rules. This is all the guidance we can have, all that we have ever had, and all that we need.

*Why Do We Need to Incorporate Moral Psychology?*

The answer to the question of why moral theory needs a robust moral psychology is this: our morality is a human morality, and it must thus be a morality directed to our human concerns, realizable by human creatures like ourselves, and applicable to the kinds of problematic situations we encounter in our lives. This means that we cannot do good moral theory without knowing a tremendous amount about human motivation, the nature of the self, the nature of human concepts, how our reason works, how we are socially constituted, and a host of other facts about who we are and how the mind operates. Moreover, we cannot know how best to act unless we know something about the details of mental activity, such as how concepts are formed, what their structure is, what constrains our inferences, what limits there are on how we understand a given situation, how we frame moral problems, and so forth. Without knowledge of this sort, we are condemned to either a fool's or a tyrant's morality. We will be fools insofar as we make stupid mistakes because we lack knowledge of the mind, motivation, meaning, communication, and so forth. Or we will suffer the tyrannical morality of absolute standards that we impose on ourselves and others, without any attention to whether people could actually live up to such standards, apply them to real situations, and improve life by means of them.

Over seventy years ago, John Dewey made the case for the empirical character of moral theory, arguing that because morality involves deliberation about possible courses of action, a vast range of empirical knowledge about action, desire, and reasoning is centrally relevant to moral philosophy:

But in fact morals is the most humane of all subjects. It is that which is closest to human nature; it is ineradicably empirical, not theological nor metaphysical nor mathematical. Since it directly concerns human nature, everything that can be known of the human mind and body in physiology, medicine, anthropology, and psychology is pertinent to moral inquiry. . . . Hence physics, chemistry, history, statistics, engineering science, are a part of disciplined moral knowledge sofar as they enable us to understand the conditions and agencies through which man lives, and on account of which he forms and executes his plans. Moral science is not something with a separate

province. It is physical, biological and historic knowledge placed in a human context where it will illuminate and guide the activities of men.<sup>4</sup>

I will argue that the exclusion from moral philosophy of a robust moral psychology stands directly in the way of genuine moral understanding and insight. Once we challenge this foundational dichotomy, it becomes necessary to reevaluate the nature and purpose of moral philosophy. The bottom line is that moral philosophy should be a theory of moral understanding, which necessarily incorporates the empirical results coming from studies in moral psychology and the cognitive sciences.

A comprehensive moral psychology would include at least the following types of inquiry:

1. *Personal identity.* What are the cognitive, affective, and social dimensions of the formative process by which a person develops an evolving sense of self-identity?
2. *Human ends and motivation.* What is the structure of motivation by which goals and purposes develop? Where do our "ends" come from? What gives rise to our conception of various goods?
3. *Moral development.* Are there stages through which people normally pass as they develop what we regard as a mature moral consciousness? Are these stages universal, or do they vary according to race, gender, or cultural differences?
4. *Conceptualization.* What is the semantic structure of human concepts? Where do our concepts come from, and how are they extended to cover new experiences? Are concepts defined by necessary and sufficient conditions, or do they have a more open, variegated, and imaginative internal structure?
5. *Reasoning.* What is the nature of moral deliberation? Is it deductive, inductive, or perhaps an imaginative exploration of possibilities for concrete action? Is it constrained in any way, or is it radically subjective and relativistic?
6. *Affect.* What is emotion? Is affect separable from conceptualization, reflection, and reasoning, or is it inextricably woven into the fabric of all experience? How is affect related to motivation? Is there cognitive and inferential structure to the emotional dimensions of experience?

This is only a partial list of the basic areas within moral psychology and the cognitive sciences that must be considered part of any adequate moral theory. Most people do not think of cognitive science as having any bearing on morality. They hold this prejudice for two basic reasons. First, they are biased against the "empirical," since they hold some version of

the fact versus value dichotomy. Second, they have an extremely narrow conception of cognitive science as formalist, reductionist, and inhumane.

But the fact is that the cognitive sciences have evolved considerably in the last decade. A second generation of cognitive science has emerged that is neither reductionist nor overly formalist. Traditional, first-generation cognitive science was defined by artificial intelligence, information processing psychology, generative linguistics, and formal model theory. It took the MIND AS COMPUTER PROGRAM metaphor quite seriously, and it had (and has) virtually nothing to say about morality, politics, social theory, and social relations. By contrast, the newly emerging second generation of cognitive science recognizes the embodied and imaginative character of all human conceptualization and reasoning. It focuses on the social, interactive, evolutionary character of human experience and understanding. It looks at human cognition as embodied in a growing biological organism that is interacting and co-evolving with its physical, social, and moral environments. Cognitive science of this sort has plenty to say about morality, and it has a major contribution to make to moral understanding.

It is obviously impossible to survey the full range of relevant empirical results from the cognitive sciences that bear directly on moral theory. I propose to consider just one small part of the new discoveries we are making in second-generation cognitive science that change our view of what moral theory is. In particular, I call attention to the metaphoric nature of our most basic moral concepts and then ask whether this fact requires us to reassess both our understanding of moral experience and our conception of moral theory.

### *The Metaphoric Nature of Moral Understanding*

In the past several years, one of the most robust and potentially revolutionary findings about the mind has been the discovery that the human conceptual system is fundamentally and irreducibly metaphoric. A large and rapidly growing number of studies have shown that our basic concepts in virtually every aspect of human experience are defined by systems of metaphors.<sup>5</sup> A conceptual metaphor is a mapping of conceptual structure from a source domain, which is typically some aspect of our concrete bodily experience, onto a more abstract or less highly articulated target domain. It is crucial to keep in mind that conceptual metaphors are conceptual. They are structures in our conceptual system, not merely propositions or linguistic entities. They involve conceptual structure, the basis for the inferences we draw from the metaphor. The content and logic of the source domain thus determines our understanding of the target domain. In other words, the reasoning we do about the target domain is based on the embodied corporeal logic of the source domain. In this way, our

systematic conceptual metaphors do not merely highlight preexisting structures in two different domains; rather, the structure and knowledge pertaining to the source domain partly construct our knowledge in the target domain.

As an example of the value and importance of moral psychology for moral theory, I focus on some recent work on the metaphorical nature of our basic moral concepts.<sup>6</sup> We are just beginning to examine the complex web of systematic metaphors by means of which we define our values, ends, actions, principles, and every other aspect of our moral experience. Moreover, because our moral concepts are defined by systems of metaphors, our moral reasoning is based on the logic of these metaphors.

An incident in Amy Tan's *The Joy Luck Club* provides a concrete example of what I mean by metaphors of morality. Ying-ying was a beautiful, refined young woman living a luxurious and carefree life with her wealthy family just before World War II. "When I was a young girl in Wushi," she remembers. "I was *lihai*. Wild and stubborn. I wore a smirk on my face. Too good to listen. I was small and pretty. I had tiny feet which made me very vain. . . . I often unravelled my hair and wore it loose." At sixteen she finds herself inexplicably attracted to an older man from another town. Within six months she is married and then realizes that she has actually come to love him.

No sooner is she married than she realizes that he is a womanizing drunkard. He abuses her emotionally, impregnates her, pursues a series of extramarital affairs, and eventually abandons her for an opera singer. This public infidelity humiliates, shames, and destroys her. He has taken her soul and left her a mere ghost:

So I will tell Lena of my shame. That I was rich and pretty. I was too good for any one man. That I became abandoned goods. I will tell her that at eighteen the prettiness drained from my cheeks. That I thought of throwing myself in the lake like the other ladies of shame. And I will tell her of the baby I killed because I came to hate this man.

I took this baby from my womb before it could be born. This was not a bad thing to do in China back then, to kill a baby before it is born. But even then, I thought it was bad, because my body flowed with terrible revenge as the juices of this man's firstborn son poured from me.

How are we to understand the logic by which this tortured innocent comes to kill her baby? In her mind it is an act of revenge. But what is the logic of revenge? In brief, Ying-ying's husband has taken her most precious possession: her spirit, her *chi*, her honor. She has "lost face" and is in shame. Ying-ying exacts her revenge by taking the most precious posses-



sion she can from him: his firstborn son: "When the nurses asked what they should do with the lifeless baby, I hurled a newspaper at them and said to wrap it like a fish and throw it in the lake." She symbolically drowns the baby in the lake, just as the women of shame drown themselves in the lake.

The logic of this tragic action stems from what I have called elsewhere the MORAL ACCOUNTING metaphor, which is concerned primarily with what we owe to other people to increase their well-being and what they, in turn, owe to us. Basically, we understand our moral interactions metaphorically as a species of economic transaction, according to the following conceptual mapping:

*The MORAL ACCOUNTING Metaphor*

<i>Commodity Transaction</i>	<i>Moral Interaction</i>
Objects, commodities	Deeds (actions), states
Utility or value of objects	Moral worth of actions
Wealth	Well-being
Accumulation of goods	Increase in well-being
Causing increase in goods	Moral = causing increase in well-being
Causing decrease in goods	Immoral = causing decrease in well-being
Giving/taking money	Performing moral/immoral deeds
Account of transactions	Moral account
Balance of account	Moral balance of deeds
Debt	Moral debt = owing something good to another
Credit	Moral credit = others owe you something good
Fair exchange/payment	Justice

This conceptual mapping provides an experiential basis for a large number of inferences that we draw in evaluating ethical conduct. We use our basic knowledge of the source domain (economic transactions) to make moral inferences about situations in the target domain of moral interactions. Consider, for example, our knowledge about wealth and how it generates inferences about morality via the WELL-BEING IS WEALTH metaphor. Wealth is something that one amasses by owning property (land, commodities) or its surrogate, money. Typically, wealth is a product of labor,

which is to say that people earn it by their work, although it may come to them by other means, such as inheritance. Being wealthy usually makes it possible for people to have more of the things they want to have and to do more of the things they want to do. It allows them to satisfy their needs and desires, and it may enhance the quality of their existence. There is a limited amount of wealth available in the world, and it must be divided up among many people. Fair exchange gives each person what is due him or her.

In the context of the MORAL ACCOUNTING metaphor, we thus come to understand moral well-being as wealth, according to the following mapping:

*The WELL-BEING IS WEALTH Metaphor*

*Financial Domain*

Wealth

Accumulation of goods

Profitable = causing increase in wealth

Unprofitable = causing decrease in wealth

*Moral Domain*

Well-being

Increase in well-being

Moral = causing increase in well-being

Immoral = causing decrease in well-being

The WEALTH metaphor is one of the two or three most important conceptions of moral well-being that we have. It shows itself in the ways we think and talk about well-being—for example:

She has had an undeservedly *rich* life.

The cynics of the world lead *impoverished* lives.

Doing disaster relief work has *enriched* Sarah's life immeasurably.

Prince Charles *profited* from his relationship with Princess Di. He is certainly a better person now.

I've had a *wealth* of happiness in my life.

Nothing can compare to the *riches* of family, friends, and loved ones.

Within the MORAL ACCOUNTING metaphor, the WELL-BEING IS WEALTH metaphor gives rise to definite inferences about our moral obligations. That is, we reason on the basis of the metaphor. Given the conceptual mapping from source to target domain and based on our knowledge of the source domain, we then develop a corresponding knowledge in the target domain and draw the appropriate inferences. Moral well-being comes to a person as a result of his or her own efforts and also as something given by the good actions of other people. Moral well-being is something that can accumulate and that can also diminish. The more well-being you have, the better off you are. Immoral action decreases well-being. Consequently,

moral acts toward others (acts that increase their well-being) put them in moral debt to you and thereby give you moral credit; you deserve an equal amount of well-being in return for what you have given.

The MORAL ACCOUNTING metaphor thus gives rise to a pattern of reasoning about our duties, rights, and obligations. On the basis of this metaphor, we reason about what is fair, and our moral discourse reveals this underlying conceptual metaphor system. If you do something to diminish my well-being, then you incur a debt to me, morally speaking, since you are expected to “pay me back” for what you have taken. When you perform noble acts, you build up moral credit. Thus, we say such things as:

*In return* for our kindness, she *gave us* nothing but grief.  
 In judging him, *take into account* all the good things he has done.  
 I’m holding you *accountable* for her suffering.  
 When you compare his kindness with what he is accused of doing, it just *doesn’t add up*.  
 All her sacrifices for others surely *balance out* the bad things she did.  
 His noble deeds *far outweigh* his sins.  
 Mary certainly *deserves credit* for her exemplary acts.  
 I *owe* you my life!  
 I couldn’t possibly *repay* your kindness.  
 Milken *owes a great debt* to society for his evil doings.  
 You must *pay* for your selfishness.

Elsewhere I have shown how the MORAL ACCOUNTING metaphor gives rise to a set of at least five basic schemas that people use to evaluate the moral merit of various actions and to determine what is due them, as well as what they owe others. Sarah Taub has outlined schemas for reciprocity, retribution, revenge, restitution, and altruism by which we draw inferences from the MORAL ACCOUNTING metaphor in deciding who gets moral credit. Take, for example, the REVENGE schema. Let us say that you do something bad to me and thereby diminish my well-being. In this sense, you have taken something from me—some of my well-being—and, via MORAL ACCOUNTING, you now owe me something that will increase my well-being to compensate for what you have taken away. But you will not give me back the measure of well-being that you owe me. Therefore, *I* balance the moral well-being books by taking an equal measure of *your* well-being, thus diminishing your moral wealth. That is why we speak of taking revenge on someone; we are taking something good from the person. The REVENGE schema thus has the following structure:

*The REVENGE Schema*

*Event:* A gives (does) something bad to B.

*Judgment:* A owes something good to B.

*Complication:* A will not give *something good* to B.

*Expectation:* B should take *something good* from A.

*Moral inferences:* A has an obligation to give *something good* to B.

B has a right to receive *something good* from A.

*Monetary inference:* B exacts payment from A.

For example:

Revenge is "an eye for an eye."

Carry *took revenge* on her classmates.

"I'll *make you pay* for what you did!"

"I'll *take it out of your hide*."

"He'll *get even* with you for this."

"Jane *owes you one* for that." (What she owes is something bad that diminishes your well-being.)

We are now in a position to see why Ying-ying does what she does. We can understand the logic of her reasoning that is based on the REVENGE schema for interpreting the MORAL ACCOUNTING metaphor. In addition to the formal structure of the MORAL ACCOUNTING metaphor, she uses two additional metaphors:

1. FACE (HONOR) IS A VALUABLE POSSESSION.
2. A CHILD IS A VALUABLE POSSESSION.

Ying-ying's husband has taken her spirit. She has lost face and is shamed. She takes away the spirit of his firstborn son. It is an empty revenge that leaves her a ghost floating through time:

I became like the ladies of the lake. I threw white clothes over the mirrors in my bedroom so I did not have to see my grief. I lost my strength, so I could not even lift my hands to place pins in my hair. And then I floated like a dead leaf on the water until I drifted out of my mother-in-law's house and back to my family home.

The REVENGE schema and the other schemas Taub has identified are all modes of expectation, evaluation, and inference that follow from the various ways in which the MORAL ACCOUNTING metaphor can be filled in by various conceptions of well-being and differing kinds of actions. They are constitutive of a large part of the moral reasoning we do when we are trying to decide what to expect from others and how we ought to treat them.

The *Joy Luck Club* example reveals importantly that there are at least two levels of conceptual metaphor operative in our moral judgments. The first ("higher") level consists of metaphors for our moral interactions generally, such as the MORAL ACCOUNTING metaphor, which sets the parameters

of our judgments about what is due us and what we owe others. Metaphors at this level define our moral framework and fundamental moral concepts. But in order for our moral frameworks to be applied to concrete situations, we need a second level of metaphor for conceptualizing the situations. The REVENGE schema, for example, is empty without the metaphors of BABY AS VALUABLE POSSESSION and FACE AS VALUABLE POSSESSION that give content to the schema and make it applicable to Ying-ying's situation. In sum, our moral reasoning typically depends on which of several possible metaphors we use at these two basic levels: (1) adopting a particular metaphorically defined framework for our interactions and (2) filling that framework in with metaphors that connect it to the particular situation (such as whether we understand the BODY AS A VALUABLE POSSESSION). These two levels must fit together to give concrete moral inferences. It follows also that moral critique can be directed at either or both of these levels, since we can criticize both our general moral framework and the more specific metaphors we use to understand aspects of situations.

### *Basic Metaphors for Morality*

MORAL ACCOUNTING is only one of several fundamental metaphors that define our moral understanding and reasoning at this first, higher level. So far, we have discovered a small number of other basic metaphor systems for various parts of our conception of morality. Although this list is by no means complete, it does set out some of the most fundamental moral concepts and shows how we reason from them. Here is the list of metaphors as we currently understand them:

1. MORAL ACCOUNTING. Our good deeds increase the moral well-being of others (via WELL-BEING IS WEALTH). They earn us moral *credit*. We *owe* others for the good things they have done for us. We should *repay* their kindness. Our evil deeds create a *debt* to other people and society in general.

*Inference patterns:* Wealth is a valuable commodity that can be amassed, earned, wasted, stolen, given away according to standards of fair exchange. Therefore, our moral interactions are regarded as modes of moral exchange in which well-being is amassed and lost and in which people build up moral credit and create moral debts through their actions.

2. MORALITY IS HEALTH/IMMORALITY IS SICKNESS. Health requires cleanliness, exercise, proper diet, and rest. When moral well-being is understood as health, it follows that all forms of moral sickness are bad. Bad deeds are *sick*. Moral *pollution* makes the soul sick. We must strive for *purity* by avoiding *dirty deeds*, moral *filth*, *corruption*, and *infection* from immoral people.

*Inference patterns:* Sick people spread disease, so we try to stay away from them and to maintain the cleanliness and bodily conditions that allow us to resist disease. Therefore, if moral evil is a disease, we must quarantine those who are immoral (by censoring them and shunning their company), so that we are not exposed to their influence. We must keep ourselves clean, pure, and protected from moral infection.

4. BEING MORAL IS BEING UPRIGHT. When we are healthy and strong, we stand upright against disease and natural forces that might knock us off our feet or lay us low. We have power and control. This same logic applies to being morally *upright*. "The Fall" (into sin and wickedness) is caused by the *force* of evil.

*Inference patterns:* Natural forces can knock you off your feet and make you lose control, thereby being unable to function successfully. Moral evil is an *ever-present* force that will cause you to lose self-control and the ability to do what is right. Morality is thus a struggle to maintain your moral uprightness, balance, and control.

There are two fundamental dimensions of being morally upright:

4A. MORAL STRENGTH. When you are morally *strong*, you have control over your *lower* self (your desires, passions, and emotions). Morality is a *struggle* between the *higher* moral self and the *forces* of our bodily selves. *Willpower* is essential to maintaining the proper *control* of our passions and *baser* desires.

*Inference patterns:* Staying in control requires strength to manage the natural forces acting upon you. In moral control, the rational will must be strong if it is to manage the powerful forces of desire and passion that lead you to pursue your animal instincts and wants.

4B. MORAL BALANCE. When you are balanced in an upright posture, you cannot be easily knocked over. When you are internally balanced, every organ works together for health and well-being. Moral *balance* is essential for moral *health*. Each part of a person must perform its proper moral function, lest he or she *fall* into evil ways.

*Inference patterns:* Keeping a balanced posture is essential for dealing with natural forces that would upset your controlled functioning in your environment. Moral balance therefore is necessary if you are going to maintain control over your ability to act as moral reason requires.

5. BEING MORAL IS BEING IN THE NORMAL PLACE. According to a pervasive metaphor for human action, which I have named the EVENT STRUCTURE metaphor, actions are self-propelled motions along paths to destinations (goals). Certain ends (destinations) are moral *ends*. They are ends we have a duty to realize through our actions. The moral path is *straight and narrow*. Moral *deviance* is a *straying* from the true path. *Violating other people's boundaries* is immoral.

*Inference patterns:* Motion along a path gets you to your desired destination. In purposeful action, ends or goals are metaphoric destinations on these motion-paths. Moral "ends" are the metaphorical places we should strive to reach, and they are socially, religiously, and morally established. Being moral is going where you ought to go, along paths set up by society. Deviance is immoral because it takes you away from or out of the region you ought to be in, along with other people. Deviants can lead other people astray, and so they are perceived as a threat to the moral community.

6. MORALITY IS OBEDIENCE. Our parents lay down for us rules of acceptable behavior, and they enforce those rules with sanctions, such as punishments and rewards. Moral *obedience* requires that we *follow rules* (obey moral laws) that tell us how to act. Reason *issues* moral *commandments* that we are obliged to obey.

*Inference patterns:* It is assumed that our parents (as authorities) have our best interests at heart and act for our benefit and that they also know what is best for us as children. Culturally, then, moral authorities are designated people who are supposed to know what is best for you and who act in the interest of your moral well-being. Obeying these authorities is the morally correct thing to do. The moral authorities can be actual people (the *pope*, religious leaders, people who are wise), or they can be metaphorical personifications (Reason, Law, the State).

7. MORAL ORDER. Having things in their proper order (in their place) is necessary for successful functioning. Society is a system—a metaphoric machine, person, building, organism—in which everything must be *in order*, if we are to live and flourish. That is why we need *law and order*. Moral *chaos* threatens to destroy society, causing widespread *breakdown, disintegration, and malfunction*.

*Inference patterns:* For every kind of system, mechanistic or organic, things must be "in order" if the system is to function properly. Things must be in working order. Morally, if everything is not "in order," society will break down and cease to function. Therefore, any disorder or chaotic activity is perceived as a serious threat to society and to human well-being.

8. MORALITY IS LIGHT/IMMORALITY IS DARKNESS. It is scary and dangerous in the dark. One can neither see nor function well in the dark. We tend to regard darkness as the harbinger of bad happenings. Moral *darkness* is a threat to our basic well-being. The *dark side* threatens to overcome the *light* in us. Evil is a *dark force*, and the *Prince of Darkness* is the most evil being of all.

*Inference patterns:* When it is dark, you cannot see things. You stumble around, lose your way, and cannot function efficiently. Metaphorically, UNDERSTANDING IS SEEING, and so the darkness of evil makes you incapable

of seeing the good and knowing what is right and wrong. You lose your moral bearings and stumble about, not being able to get where you ought to be going (to your moral ends). Moral darkness brings ill-being to us.

9. MORAL PROJECTION. Human perception is ineliminably perspectival. We always experience any object from a particular point of view, and the more perspectives we can take up on the object, the more objective our knowledge is considered to be. We rise above some of the limitations of our own way of seeing things by "seeing things as others see them." Morally, projecting ourselves into another person's way of experiencing things helps us rise above our own prejudices in order to act more humanely.

*Inference patterns:* If you want your perceptual knowledge to be as objective as possible, you try to achieve as many perspectives as possible. If you want objective moral knowledge, you must take up the *moral point of view*—the point of view of a "moral person in general." This means that you must be able to abstract from your prejudices and take up the standpoint of an ideal moral judge who decides how to act on grounds that hold for every person, not just for this or that particular person. Moral empathy is a form of putting yourself in the place of another. All moral theories founded on the basis of a rational moral agent require the metaphor of the projected self (for example, Rawls, ideal-observer, and universal standpoint theories).

This list of basic metaphors for morality is partial and needs extensive further analysis, for example, setting out the structure of the source domain for each metaphor, explaining why that particular source domain is used, laying out the conceptual mapping from source to target domain, and showing how this mapping constrains our moral reasoning. However, even this brief list suggests two very significant points about the metaphoric nature of our moral understanding. First, it is vital to notice that basic moral terms like "ought" and "should" really have meaning and lead to moral inferences only through one or more of the above metaphors, along with the second-level metaphors by which we conceptualize actual situations. "Ought" means one thing and supports certain very specific moral inferences in the context of the MORAL ACCOUNTING metaphor, compared to the quite different set of inferences that it generates relative to the MORALITY IS HEALTH metaphor. According to the MORAL ACCOUNTING metaphor, for instance, "ought" is spelled out in terms of economic transactions of fair exchange, credit, and debt. MORALITY IS HEALTH, by contrast, establishes imperatives that direct us to fight moral sickness and promote certain states of moral flourishing within individuals and the community. "Ought" therefore gets its content and concrete applicability by means of its role in metaphorically defined moral frameworks.<sup>7</sup>



Second, it is remarkable that there seem to be so few basic source domains for our metaphors of morality. Why should we have these source domains and not others? My research so far suggests the following general answer. First, these source domains appear to be universal in human experience because they depend on the nature of our bodies and their typical interactions with the types of environments we inhabit. Second, these source domains are characteristically tied up with our sense of personal and communal well-being, growth, and satisfaction, which makes them suitable source domains for metaphors of morality. It is no accident that source domains such as health, balance, strength, movement to a place, obedience, and light/dark are intimately tied up with our sense of human flourishing, and so they are prime candidates for universal metaphors of morality. Whether these are, in fact, universal source domains is a matter for empirical study, but there is some evidence to think that they may be found in all cultures, even if they are not elaborated in just the same way in each culture.

#### *Metaphor and Moral Reasoning*

The most important epistemological and moral implication of the fact that our basic moral concepts are defined by metaphors is that we reason on the basis of these metaphors about how we ought to act and what kind of person we ought to be. The logic of the source domain, as it is mapped onto the target domain, constrains the inferences we make about the target domain of morality. We have seen this already in the way that the REVENGE schema leads to judgments and actions within the framework of the MORAL ACCOUNTING metaphor. The crucial point is that each metaphor has its own logic and generates epistemic entailments about the target domain (which is here some part of morality).

In order to show how strongly these metaphors constrain our moral reasoning, let us consider the metaphor MORALITY IS HEALTH. The relevant conceptual mapping is:

#### *The MORALITY IS HEALTH Metaphor*

<i>Physical Health</i>	<i>Moral Behavior</i>
Health	Well-being
Sickness	Moral degeneration
Pollution	Cause of evil
Being diseased	Being morally depraved
Physical exercise	Moral training
Growth	Moral improvement

Notice the very strong way in which the logic of the source domain (physical health) determines the logic of the target domain and constrains our reasoning. What is it that makes you unhealthy? The answer is, any cause of disease—infection, plague, pollution, filth, and, in general, things that you find disgusting. What promotes health? The answer is, exercise, watching what you eat, avoiding those who are sick, staying clean, and so forth. Knowledge of this sort from the source domain carries directly over into our reasoning about morality. We speak of moral and social *diseases*, of *sick* acts, of *dirty* pictures that *pollute* people's minds, of *filth* and *smut*, of *disgusting* behavior, and *cancer* at the heart of society. Sick people can spread disease. Consequently, in order to avoid becoming sick ourselves, we stay away from them, quarantine them, try to kill the disease, and try to keep ourselves clean. Morally, then, we argue that immoral people can *infect* others with their evil. We do everything we can to distance ourselves and our children from the causes of immorality and the people we regard as immoral. We try to *clean up* our schools and our towns. We treat illegal drugs as a *plague* on society. We think that association and contact with someone who does an immoral act can cause a *diseased* mind.

In short, the *logic* of the metaphor determines our expectations, our reasoning, and our actions. The HEALTH metaphor is evidenced in a huge variety of expressions we use to make moral judgments—for example:

His intentions were *pure*, even if things didn't work out well.

She has no moral *blemishes*.

"... without *spot* of sin."

"O Lord, create a *pure* heart within me."

"What a *stinking, low-down, dirty* trick, you miserable rat!"

Scarlett was *washed clean* of her sin.

Pornography *pollutes* the mind and soul.

We must keep that *filth* out of our schools.

Crime is a *plague* that *infects* us all.

He lives in a cocaine *sewer*.

What a *sick, disgusting* thing to do.

These conceptual metaphors are not merely optional ways of talking about morality. There is nothing optional about them at all, and they are not merely matters of words. They are the means by which we define our moral concepts. Although we may not be limited to just one unique metaphor system for a particular concept (for example, we have both WELL-BEING IS WEALTH and WELL-BEING IS HEALTH), neither can we use just any metaphor, especially since this is seldom a matter of conscious choice and the range of metaphors available is relatively small. Most important, the nature and structure of the source domain constrains the inferences we

make about the aspect of morality that is the target domain. The metaphor, in other words, sets limits on our reasoning about how we ought to behave and how we ought to regard others. If, for instance, we understand moral well-being as WEALTH, we will act and reason quite differently than if we understood it as HEALTH. The logic of the WEALTH metaphor contains notions of fair exchange, quantification, and balance; while the logic of moral HEALTH emphasizes avoidance of immoral people, staying pure, fighting moral disease, and maintaining moral discipline.

### *How Cognitive Science Changes Ethics*

Having surveyed some of the basic metaphors that define our moral understanding, the nagging question arises, So what? "So what?" the moral apriorist will ask. "What difference could it possibly make to learn that people typically use metaphors to understand their experience? We want to know how they *ought* to reason, not how they tend to reason."

The answer to this question is clear and straightforward, and in offering an answer, I am suggesting in general how empirical studies in the cognitive sciences bear on morality and moral theory. The general answer is that our morality is a *human* morality, one that must work for people who understand, and think, and act as we do. Consequently, if moral theory is to be more than a meaningless exploration of utopian ideals, it must be grounded in human psychology.

The moral purist, in pursuit of the illusory ideal of a strict governance theory of morality, demands a nonexistent direct connection between moral understanding and morally correct action. The only answer that a moral purist will allow is one that shows how learning about the metaphoric structure of morality, for instance, would lead, in a step-by-step fashion, directly to rules that would tell us how to act. But this is not possible in any but the most obvious, well-worn, unproblematic cases. Knowing the nature and entailments of the MORAL ACCOUNTING metaphor does not tell us whether MORAL ACCOUNTING is a good form of moral interaction in any particular situation. However, knowing all we can about the MORAL ACCOUNTING metaphor can help us make informed judgments about the probable consequences of acting on the basis of this particular metaphor.

It is extremely important to see that the moral purist's charge that cognitive science and moral psychology have no direct bearing on moral theory is based on an illusory fact-value dichotomy that manifests itself in two main fallacies:

1. *The Independence of Facts fallacy:* Facts are independent of any value assigned to them.
2. *The Irrelevancy of the Empirical fallacy:* Facts do not tell you what value to assign to them.

The Independence thesis has been demolished most prominently by arguments in the philosophy of science showing that whatever is counted as a "fact" depends on certain values we have, such as our interests, purposes, criteria of importance, or models. The Irrelevancy thesis is more difficult to falsify because it contains the kernel of truth that if we understand moral psychology as merely a description of why this or that person or community reasons or acts in a certain way, then indeed this will not tell you how they ought to reason and act. These kinds of descriptions do not supply a basis for moral critique.

Nevertheless, the Irrelevancy thesis can be shown to be false once we introduce a much more profound sense of moral psychology and cognitive science that does have normative implications for our moral judgments. As it applies to moral theory, cognitive science in this richer sense is the empirical study of how we conceptualize values and reason about them. This aspect of moral psychology *does* have normative and critical implications, because it gives us insight into the nature of our values and how they constrain our inferences about moral matters. For example, the empirical study of our moral conceptual system reveals the metaphors that define our moral frameworks, and it can open our eyes to the limitations of this or that metaphor of morality. It can show us what our metaphors highlight and what they hide. It reveals the partial nature of any metaphorical conceptualization and of the reasoning we do based on each metaphor, and it shows us that we may need multiple conceptualizations to discern the full range of possibilities open to us in a given situation. Knowledge of this sort is knowledge that should influence our judgments and actions. It is knowledge that comes from what I earlier called the psychology of moral understanding, which I contrasted with trivialized moral psychology, that is, moral psychology with blinders. A rich psychology of moral understanding looks not merely at people's beliefs and motivations but especially at their deepest moral concepts and the reasoning that stems from them.

The absence of rules for deriving moral judgments from knowledge about metaphorical concepts is not something to be lamented. It is simply a fact about the complexity of human moral understanding, and it is an extremely important fact that has the following significant implications for ethics.

#### *Conceptual Analysis*

If our basic moral concepts are metaphoric, then conceptual analysis must presuppose some view of the nature of metaphor as underlying the analyses. Any moral theory that does not recognize the metaphoric nature of moral concepts must be inadequate, and probably disastrously so. Furthermore, the adequacy of the theory will depend on the adequacy of its

theory of metaphor. If we are going to get insight into morality, we need a view of metaphor that recognizes its central role in understanding and reasoning. Whatever remains of "metaethics," therefore, is to a significant extent an exercise in metaphor analysis.

#### *Moral Reasoning*

If our basic moral concepts are metaphoric and if we use metaphors to frame the situations we are deliberating about, then our moral reasoning is primarily an exploration of the entailments of the metaphors we live by. For the most part, then, moral reasoning is not deductive, and it is not primarily a matter of applying universal moral principles or rules to concrete situations. I have shown why this model cannot work for the kinds of beings we are and for the kinds of situations we encounter. The reason is that the traditional deductive model has no place for metaphoric concepts, or for any concepts that do not have classical (that is, necessary and sufficient conditions) structure.<sup>8</sup>

#### *Partial Understanding*

It follows from the imaginative nature of moral concepts and reasoning that no understanding is exhaustive or comprehensive. Human moral understanding is a complex cluster of metaphor systems, some of which are mutually inconsistent, and yet we manage to live with them and plot our lives by them.

#### *Beyond Absolutism*

Because our moral understanding is necessarily partial, morality is not a set of absolute, universal rules but an ongoing experimental process. We must continually be experimenting with new possibilities for action, new conceptions of human flourishing, and new forms of interaction that permit us to adjust to, and also to manage, the ever-changing conditions of human existence. As long as we and our entire ecological situation are evolving, morality must remain experimental. Any attempt to codify this procedure into a final method or absolute principles is a recipe for moral rigidity and obtuseness.

#### *Grounded Moral Theory*

The partial, nonabsolute character of our moral understanding might make one think that morality is historically and culturally contingent in a radical way. It would then seem that there is no point in trying to construct a normative moral theory. This is a mistaken view. The most basic source domains for our metaphors for morality are grounded in the nature of our bodily experiences and tied to the kinds of experiences that make it possible for us to survive and flourish, first as infants and then as developing

moral agents. Whether these basic source domains are universal is an open question that awaits further cross-cultural investigation. But if anything is universal, we have good reason to think that structures such as these will be. I believe that these experiential source domains provide general constraints on what can be a psychologically realistic morality, as well as an adequate moral theory. The general nature of such constraints suggests, as Owen Flanagan has argued at length, that there will always be a plurality of appropriate conceptions of human flourishing and a range of possible ways of realizing such conceptions of the good. Although these constraints do not underwrite a universal governance theory, they do limit the range of acceptable alternatives.

### *Moral Imagination*

Moral deliberation is an imaginative enterprise in which we explore the possibilities for enhancing the quality of human existence in the face of current and anticipated conditions and problems. When we are trying to figure out the best thing to do in a given situation, we are tracing out the implications of various metaphors to see what they entail concerning how we should act. Projecting possible actions to determine their probable results, taking up the part of other people who may be affected, and reading with sensitivity the relevant dimensions of a particular situation are all forms of imaginative activity.

People who stress the imaginative and affective dimensions of human understanding are often mistakenly accused of being irrationalist and subjectivist. This serious misinterpretation is the result of a continued adherence to traditional rigid distinctions between such capacities as perception, imagination, feeling, and reason. Stressing the imaginative nature of our moral understanding in no way impugns the rationality of morality. I am arguing here for an enriched conception of human reason as fundamentally imaginative. My point is that moral reasoning is a much richer, more complex, and more flexible capacity than it has been conceived to be in traditional Enlightenment accounts of practical reason. Moral reasoning *is* reasoning, but it is a reasoning that is thoroughly imaginative in character.

### *What Should a Theory of Morality Be?*

A theory of morality should be a theory of moral understanding. Its goal should be moral insight and the guidance and direction that come from a deep and rich understanding of oneself, other people, and the complexities of human existence. At the heart of moral reasoning is our capacity to frame and to realize more comprehensive and inclusive ends that make it possible for us to live well together with others. It involves an expansive

form of imaginative reason that is flexible enough to manage our changing experience and to meet new contingencies intelligently. The key to moral intelligence is to grasp imaginatively the possibilities for action in a given situation and to discern which one is most likely to enhance meaning and well-being.

The idea of moral theory as providing governance through rules and principles is fundamentally mistaken. In fact, it is counterproductive to the extent that it overlooks the changing character of experience and does not allow us to see creatively new possibilities for action and response. As Dewey saw, what moral principles we have are not technical rules but rather "empirical generalizations from the ways in which previous judgments of conduct have practically worked out."<sup>9</sup> They are summaries of strategies that have proved more or less useful for the kinds of situations previously encountered. But they must never be allowed to solidify into absolute rules, for then the opportunity for moral growth and progress is undermined.

It should now be obvious why I think that the alleged split between moral theory and moral psychology is not just bogus but detrimental to a sound moral philosophy. The goal of moral psychology and moral philosophy alike should be understanding and liberation. Moral philosophy will give us the guidance that comes from moral understanding, critical intelligence, and the cultivation of moral imagination. It will not tell us what to do, but it will help us struggle to discern better from worse possibilities within a given situation. Moral philosophy cannot, and never did, give us an adequate theory of moral governance. Once we are liberated from this illusion, we can interpret Kant's dictum—always to think for yourself—as a call for a mature attitude of continual, well-informed, critical, and imaginative moral experimentation. And in our ongoing communal moral experimentation, good cognitive science, coupled with the cultivation of moral imagination, should lead the way.

### Notes

1. Alan Donagan, *The Theory of Morality* (Chicago: University of Chicago Press, 1977), pp. 1, 7.
2. Owen Flanagan, *Varieties of Moral Personality* (Cambridge, Mass.: Harvard University Press, 1991), p. 32.
3. Samuel Scheffler, *Human Morality* (Oxford: Oxford University Press, 1992, p. 8).
4. John Dewey, *Human Nature and Conduct* (1922), in Jo Ann Boydston, ed., *The Middle Works of John Dewey, 1899–1924* (Carbondale: Southern Illinois University Press, 1988), pp. 204–205.
5. A large number of these studies of the metaphoric nature of human conceptual systems are discussed in George Lakoff, *Women, Fire, and Dangerous Things: What Categories Reveal about the Mind* (Chicago: University of Chicago Press, 1987), and in Mark Johnson, *The Body in the Mind: The Bodily Basis of Meaning, Imagination, and Reason* (Chicago: University of Chicago Press, 1987).

6. Most of the relevant work on the metaphoric nature of our moral concepts is summarized in Mark Johnson, *Moral Imagination: Implications of Cognitive Science for Ethics* (Chicago: University of Chicago Press, 1993), esp. chap. 2.
7. In *ibid.*, chap. 3, I have shown how such abstract and formal moral principles as the various formulations of Kant's categorical imperative are actually based on different conceptual metaphors, without which there could be no application of moral laws to concrete situations.
8. *Ibid.*
9. Dewey, *Human Nature*, p. 165.