INTRODUCTION

Raphael Demos, in his seminal article "On Lying to Oneself," describes self-deception in the following way: "Self-deception exists, I will say, when a person lies to himself, that is to say, persuades himself to believe what he knows is not so. In short, self-deception entails that B believes both P and not-P at the same time."[1] It is clear that Demos is here confining himself to one significant dimension—for philosophers of his bent, possibly the most significant dimension, of self-deception. Although many subsequent commentators (including Brian McLaughlin in this volume) have joined the discussion, assuming that this paradox encapsulates the general notion of self-deception, it is feared by others that this description might not capture the full complexity of the phenomenon and its ramifications. It also entails certain presuppositions about human agency that are not commonly subscribed to in all cultures, or even by all philosophers within our own tradition—that is, it entails a unitary conception of self as both subject and object of "self-deception." In fact, it is to underscore the contingent status of "the self" as a superordinated agent that this volume has been entitled *Self and Deception* rather than *Self-Deception*.

Disagreements about the phenomenon of self-deception are many and varied. They spill over into the diverse areas of concern expressed by our authors: the nature of consciousness, conditions of intentionality, rationality, gradations of rationality and irrationality, the importance of cognition in psychological processes, cognitive relativism, categories of belief, epistemic foundations of self-deception, incontinence, wish fulfillment, authenticity, hypocrisy, character modification, the nature of intentional explanation, the social construction of self and self-deception, rhetorics underlying self-deception, the positive function of irony, and so on. These domains of inquiry, most of which have already received considerable scholarly attention, attest to the breadth and the complexity of the issues surrounding self-deception.

The paradoxical nature of self-deception as a starting point is an issue that has generated a great deal of discussion and continues to do so. For

example, Mary Haight argues that self-deception is literally a paradox, and hence it cannot occur.[2] Her argument, familiar in the more analytic literature, is that deception requires a split between deceiver and deceived and hence it cannot take place in relation to one and the same person. While she focuses on what she thinks is the paradoxicality of self-deception, other exegetes such as Herbert Fingarette have attempted to provide a nonparadoxical and morally unforgiving reading of self-deception.[3]

The psychological and moral dimensions of self-deception have generated much interesting discussion. Jean-Paul Sartre's concept of 'bad faith' (*mauvaise foi*) served to bring the diverse issues surrounding self-deception to a much wider audience. The epistemic implications of being self-deceived have also attracted their share of scholarly attention.

There is also a cross-cultural dimension to this problem of self-deception, brought into focus in the opening pages of Eliot Deutsch's essay. Is it the case that the same phenomenon of self-deception is conceptualized and valued differently in different cultures, or is it—self-deception—fundamentally a different phenomenon? Is self-deception a cross-cultural issue, or is it (like alienation or *amaeru*) a culturally specific phenomenon? If it is present in some form and with some value in alternative cultures, what does it entail? To what extent do culturally specific conceptions of 'self' enter into the discussion? For example, if the peculiarly Western "mind-body" problem is entailed in the discussion of self-deception through assumptions about self, thereby "psychologizing" self-deception, what are the non-Western alternatives? Does it make sense to think of a somatic aspect of self-deception? This, we would hope, would enable us to wrestle with two questions: How can self-deception take place? And how does self-deception take place in an enlarged discursive setting?

The concept of 'self-deception,' impinging as it does on such important areas of exploration as the nature of belief, rationality, and cognition, underlines the importance of drawing on, if only for purposes of heuristics, rapidly developing fields such as cognitive science, cultural anthropology, theoretical biology, sociology of knowledge, and decision science. The interplay of the dynamics of cognition and cultural constructions of belief might well open up newer and more fruitful avenues of inquiry into cognitive relativism and motivated irrationality.

Questions of rationality and belief are at the center of current discussions of self-deception. This, of course, means that questions of bounded rationality, pseudorationality and irrationality, unbelief, and desire are pivotal in these discussions. On the one hand, there are those who argue that self-deception represents the power of irrationality in human affairs. On the other hand, some maintain that deception of self does not constitute a form

of irrationality based on incoherence of belief, but rather an incongruity between a person's self-conception and one's praxis. Fingarette advocates the latter view. Jon Elster would argue against the assumption that the function of rationality is to tell the agent what to do, so that if one behaves in a way that is contradictory, one is irrational. Instead, he would assert that rationality can do no more than exclude certain alternatives, while providing no guide to one's choice among those that remain.

The concept of 'self,' understandably enough, is central to any sustained inquiry into self-deception, the pertinent question being what sort of self is victim (or beneficiary?) of self-deception? Basing their thinking on the model of "other-deception," some would bring in notions of 'double selves,' 'multiple selves,' and 'subsystems of the self,' to address this troubling problem. Others would argue that other-deception is not an adequate or a reliable model to guide our thinking on this issue.

Similarly, the concept of 'emotionality' is important. What is the relationship between desire and belief, between emotion and self-deception? Does self-deception, in the domain of emotions, imply an agent who is misled by his or her emotions, or are there alternate ways of formulating these relationships? What then is the nature of the relationship between desire and belief?

It is now increasingly remarked upon that our investigations into the phenomenon of self-deception are excessively intrapersonal and that issues pertaining to the social dimensions of self-deception—namely, rules of human interaction and social conventions—are receiving inadequate attention. The whole question of the relationship of intersubjectivity to acts of self-deception needs to be explored afresh. Since the complex and intriguing ways in which people make sense of their behavior, construct systems of meaning, and develop strategies of interpretation are all grounded in their sociality, this interpersonal aspect requires deeper study if we are to uncover the social bases of self-deception. The fundamentally relational definition of person in the East Asian cultures might be a source of direction and encouragement in this respect.

Let us, for example, consider the concept of 'rationality.' In view of the fact that rationality is subject to public scrutiny, it has the effect of bringing into social purview the various processes that go to form rationality. Here again, the social basis of rationality as formulated and manifested in Asian cultures has the potential to promote profitable discussion.

We normally regard self-deception as a moral failure and valorize it negatively. Sartre, for instance, thought of it as the unparalleled vice. But there are others who wish to invest it with positive value, seeing it as a potential strategy of coping which serves to enhance self-respect and

3

strengthen community bonds. Taking a cross-cultural perspective on this issue will enable us to locate this problematic in a wider discursive field.

What kind of person is capable of self-deception? This question takes on a further layer of complexity when we introduce Asian traditions of thought. For example, the Hindus, most notably the Advaita Vedāntins, believe that the true Self (*ātman*) is pure, undifferentiated, self-luminous consciousness. It transcends normal consciousness, being aware of only the Oneness of being. The individual human being is regarded as representing an admixture of reality and appearance. According to early Buddhism, the self is unreal and impermanent. As everything is in a state of flux, there cannot be an unchanging self. The Buddha conceived of the self as a cluster of aggregates or *skandhas.* Confucianism conceptualizes the self as an open system which is vitally connected to the family, community, state, and cosmos. It is irreducibly interpersonal. The ideas of relationality and context are central to a robust understanding of the Confucian notion of 'self.'

These three culturally specific formulations of self are developed in the pages that follow as a necessary foundation for examining the phenomenon of self-deception. When discussing the concept of 'self,' the distinction between the private self and the public self is indeed important in some cultures. In Japanese society, for example, this binarism has deep implications in terms of action and behavior. Japanese modes of communication are closely linked to this distinction. When examining the antecedents, dynamics, consequences, and rationalizations of self-deception in Japan, one must pay careful attention to divisions such as the private and the public selves. In the Chinese tradition, by contrast, any severe distinction between public and private selves is entirely suspect.

One reason why the whole phenomenon of self-deception and the attempts to deconstruct it becomes so intriguing is because in Western philosophical discourse, the paradigms for the explication of belief and action are largely based on rational behavior. Consequently, when a seemingly irrational form of behavior such as self-deception becomes the focus of discussion, the immediate tendency is to seek to accommodate it within existing paradigms of explication—that is, to explain it rationally. Even to describe it as "irrational" organizes it in some degree. How have the less rationalistic Asian traditions grappled with this same issue, and in what ways can these efforts illuminate Western thinking?

In Eastern intellectual traditions there are a number of concepts that are centrally related to the idea of deception: 'ignorance' (*avidyā*), 'illusion' (*māyā*), 'delusion' (*moha*), 'obscuration' (*pi* 蔽), 'acting unnaturally' (*youwei* 有爲), and so on. These are closely linked to questions of knowledge, action, behavior, and emancipation. An examination of these issues opens

4

up lines of inquiry that can both complement and challenge the work carried out by Western thinkers.

In classical Greek philosophy, the idea of *akrasia* or weakness of will is closely associated with the phenomenon of self-deception. The opposite of *akrasia* is *enkrateia,* or self-control. If *akrasia* is manifested in behavior that contradicts one's better judgment, *enkrateia* is manifested in behavior that is in accordance with one's better judgement. Self-control is indeed a virtue that is highly prized in Eastern traditions of thought, whether it be Hinduism, Buddhism, or Confucianism. What light do the functional equivalents of *akrasia* and *enkrateia* as found in these traditions—darkness and clarity of heart and mind in Confucianism, for example—throw on our understanding of self-deception?

The semiotics of self-deception is another area that can be of great value in promoting profitable discussion. The ways in which the concept of 'self-deception' is articulated in everyday language, and the diverse locutions characteristic of different cultures, enable us to approach this concept from alternate vantage points. Some believe that although the locutions *self-deception* and *self-instruction* display a semiotic kinship, they are in point of fact so different phenomenologically that they should not be placed in the same category. This is the situation encountered within the context of the English language. What are some of the semiotic problems found in Chinese or Japanese languages, and what do they tell us about the conceptualization of self-deception? What is the cluster of locutions which needs to be scrutinized and reconstructed in order to pursue the discussion within these cultures?

Clearly, some very exciting work has been done in the understanding of self-deception in Western philosophy. The work of Mary Haight, Herbert Fingarette, Donald Davidson, David Pears, Amélie Rorty, and Jon Elster, as well as that of Jean-Paul Sartre, to name but a few, has opened up promising avenues of inquiry. Several of the papers in this volume take their lead from the published work of these scholars and provide us with a framework from which we can extend the discussion to the non-Western traditions. In so doing, we can expose ourselves to different styles of reasoning and different conceptual schemes characteristic of different cultures and provide a basis for Western and Eastern traditions to interrogate each other.

Brian P. McLaughlin presents an important introduction to the analytic discussion of self-deception by laying out the terrain of the current debate. The volume begins with his reflections, since he belongs to a conversation that has held important sway in Western literature on self-deception. He initiates a dialog between the "nonskeptic" who asserts that self-deception is not only possible, but pervasive (a position joined by almost all of the

contributors to this volume), and the "skeptic" who asserts that there are "excluders" of its possibility (where the term *excluders* is defined as conditions required for self-deception that cannot obtain).

There are two commonly cited excluders of self-deception, the first being *the contradictory belief condition:* the condition of X believing P and not-P at the same time. One common strategy for nonskeptics who would claim that this is only an apparent excluder is to introduce some notion of 'separation' or 'dissociation': not-P might be repressed in some way while one entertains P. Another response for the nonskeptic is to deny that self-deception entails believing P and not believing P at the same time—self-deception can be otherwise understood.

The second commonly encountered excluder is *the knowing dupe condition:* the condition of being duped by a stratagem executed on oneself, by oneself. This situation requires that both the self-deceiver and the self-deceived know that it is a stratagem, and yet the self-deceived is still duped by it.

The nonskeptic again has two avenues of response. Most nonskeptics, according to McLaughlin, do not attempt to argue that one can be taken in by a stratagem one regards as such: that is, that one can be misled by information that one knows to be misleading. Rather, they choose to argue that self-deception does not require this knowing-dupe condition at all.

McLaughlin then turns to the attempts on the part of the nonskeptics to argue that these two excluders—the contradictory-belief and the knowing-dupe conditions—are only *apparent* rather than *genuine* excluders. The nonskeptic is able to provide an example of a person with normal memory loss deceiving herself into missing an appointment she does not want to attend by misscheduling it. Both skeptic and nonskeptic can agree that this "appointment" case requires neither of these excluder conditions, and also that self-induced deception is possible. The skeptic, however, allowing that self-induced deception is necessary for self-deception, will not allow that it suffices for self-deception. The argument here is that both skeptics and nonskeptics can generally agree self-deception that P entails at some level the recognition that the weight of evidence is against P. No such requirement is necessary for self-induced deception. The woman in the appointment case never recognized at some level that the weight of evidence was against her having an appointment.

The tension between skeptic and nonskeptic lies in the following. The skeptic will insist that in a genuine case of self-deception that P, recognition at some level that the weight of evidence is against P, will give rise to the belief that not-P. Nonskeptics generally, while not conceding this claim, will allow that the self-deceiver that P might harbor some anxiety

that not-*P*. The nonskeptic will argue that the self-deceiver that *P* is able to maintain this situation by either repressing evidence of not-*P* or dissociating oneself from such evidence in some way. The problem with homuncular models of self-deception—models that depend on relatively autonomous subsystems—is that in explaining self-deception, they explain it away. Dissociation makes the phenomenon disappear. A distinction is introduced that separates deceiver from deceived that dissolves the possibility of self-deception. These homuncular models reduce to exotic cases of other-deception: something like multiple personality disorders.

In considering how separate mental divisions must be for the degree of dissociation or repression necessary to generate self-deception, McLaughlin introduces the possibility of "inaccessible" belief. Someone can believe that *P* and that not-*P*, as long as one of them is inaccessible. Beliefs, according to McLaughlin, can slip back and forth between accessibility and inaccessibility.

In explaining how the self-deceptive belief that *P* might be acquired and sustained, nonskeptics frequently appeal to rationalization, evasion, and overcompensation as activities in which self-deceivers unintentionally mislead themselves. The skeptic replies by insisting that for self-deception, one must be a deceiver, and since an unintentional misleader is not a deceiver—that is, need not perform an act of deception—there is no self-deception. Stated more simply, there is no such thing as unintentional deception. Thus the fact that unintentional misleading is possible does not make self-deception possible.

According to McLaughlin, this is the current geography of the debate. The following represents the kinds of questions which remain: Are the conditions placed on self-deception by the skeptic not simply over-rationalizing what are really only cases of unintentional misleading? This would give the nonskeptic at least a partial vindication of the notion of 'self-deception' while allowing the skeptic the claim that *literal* self-deception is impossible. The nonskeptic would reply that the unintentional misleader can be a deceiver in whatever sense a self-deceiver is a deceiver, and thus unintentional misleading suffices for self-deception.

Whatever the case, McLaughlin sees the real payoff for better understanding the architecture of our minds to lie in discovering how desire affects belief-fixing mechanisms. And this question can best be investigated by trying to determine if the kind of self-deception the skeptic has in mind is possible—that is, the toughest case of self-deception. It is only of secondary importance if there are kinds of self-deception which do not require the contradictory-belief or knowing-dupe conditions.

Annette C. Baier chooses to focus on self-deception as a kind of selective attention, a situation in which ignoring, like attending, is something

7

one must work at. Our ability to foreground and background is, far from ignorance, an important cognitive capacity that allows us a degree of impartibility in making moral judgments. Baier's essay concentrates on the line between the positive achievement of a functional and efficacious selective concentration and the negative eliding of purposely ignoring with purposeful ignorance. By understanding self-deception in terms of a motivated unwillingness or inability to recall what we have chosen to ignore, we escape the paradox that more cognitively focused accounts abjure.

Self-deception as selective ignoring is a normal aspect of the human experience, for persons, for groups, and even for nations. It can be a kind of denial, dissociation, and rationalization necessary to preserve one's self-conception intact, and as such, can be literally a life-preserving skill. Baier's interpretation of judicious self-deception as a global psychological and intellectual capacity for would-be rational human beings to deal with the fact that they are only imperfectly equipped to understand their worlds resonates comfortably with the positive turn on self-deception developed both by Amélie Rorty and by Robert C. Solomon in this same volume.

To further explore this phenomenon of self-deception, Baier goes back to Descartes in the *Meditations* where he begins by the deliberate pretence of regarding all of his previously held beliefs to be false, fortifies this pretence by introducing the possibility of the evil genius who makes the mediator the victim of deception, and then combines deliberate pretence and deception into a kind of self-deception by having the meditating *I* conspire with the evil genius to prolong the deception indefinitely. Ironically, the meditating *I* never seems to awake from what began as a deliberate strategy to avoid self-deceiving beliefs. The analysis is complex, and the real difficulty seems to lie in the failure of Descartes to stipulate explicitly when the mediator moves from pretended denials to real denials.

Baier rehearses this well-known tract to demonstrate the positive and the negative implications of self-deception. The problem of self-deceit arises when the mediator ignores the fact that what was announced as a deliberate strategy—the suspension of preconceived beliefs—is never itself suspended.

A distinctive feature of self-deceit is its tendency to confuse itself with plain deceit and to prefer the role of deceiver to deceived. It is the temporality of the *Meditations* that enables Baier to resist the importance of those analyses which resolve the paradox of self-deception by appealing to contemporaneous subsystems deceiving each other. It is the complexity of the temporal shifts in the *Meditations* between the deliberate and the inadvertent, between the deceiver and the deceived, between self-deception and deceit, that make it a worthy presentation of this phenomenon. An interesting aspect of the entire discussion of the *Meditations* that Baier reflects upon

is its seeming asocial nature. Given that there are only the relationships between mediator, demon, and God, the focus of notions like 'trust' and 'deceit' is open to speculation: Are we talking about present self in relationship to past self, partial self in relationship to another part, or the relationships that obtain among various powers within the same psyche?

It is only with the *Objections* that the meditating *I* in the *Meditations* at last has the benefit of friendly (and less friendly) counsel to demonstrate that rationally motivated self-deception in the company of our fellows is indeed difficult to sustain.

Baier concludes by observing that the more contemporary attempts to deceive ourselves philosophically by denying the phenomenon of self-deception as a way of saving our favorite rational theories contrasts rather starkly with Descartes. For Descartes takes self-deception both as his method and as his main theme and illustrates rather too clearly what self-deception is, what we risk when we engage in it, and probably most important, the price we would have to pay to give it up.

Amélie Rorty allows that self-deception can be pernicious. But the narrow understanding of Enlightenment rationality that makes self-deception the epistemological equivalent of original sin overshadows the necessary and indeed positive function of a range of self-induced illusions that sustain us in the everyday. Rorty observes that as a complex phenomenon, the negative or positive valorization of self-deception is often a matter of subjective interpretation where the same situation can be read in even contradictory ways at the pleasure of the observer. Nonetheless, there are interpretive constraints which can be invoked to bring at least some focus to a rather nebulous area in human experience. On the one hand, if we proceed from the perspective of theory construction in attempting to deal with the possibility of knowledge itself, self-deception and like concerns stay marginal to the inquiry, if they are considered at all. On the other hand, if we proceed from the perspective of everyday lived experience, leaving abstract systemic demands aside, these same theoretically marginalized concerns can be seen to be pervasive in all areas of human activity. And the dialectical nature of our philosophical tradition guarantees that generationally the discussion moves back and forth between these extremes.

Rorty underscores the social and political complexity of deception and self-deception as entailing subtle strategies of persuasion, making it exceedingly difficult to draw any sure demarcations in what is fundamentally an interactive process. Having said that, one way of moving closer to an adequate understanding of self-deception is to establish what it is *not* by clearing away popular misconceptions of it. First, self-deception is a pattern of behavioral dispositions and habits of the mind rather than isolatable incidents.

9

Second, such dispositions are socially interactive, being reinforced by encouragement frequently elicited by the self-deceiver. Third, self-deception need not involve false belief. One can deceive oneself by concentrating on true yet marginal information at the expense of considering immediately relevant data—this too is a mechanism of denial. An established disposition to ignore what is pertinent information is at once an act of self-deception and a strategy for self-deception. Further, the performance of ritually and habitually constituted patterns of behavior might entail mistaken beliefs about oneself rather self-deception.

Fourth, self-deception is as likely to be formed in the ordinary events of daily life as it is to characterize important undertakings. Fifth, self-deception has a very real social aspect, where although it is always *by* the self, even its initiation can be fostered collaboratively in one's association with others. The social and even the physical environments can encourage self-engrandizement.

Sixth, self-deception can function as a psychological habit or a disposition acquired by repetition or imitation that is really lacking in any specific desire or motivation. Finally, Rorty observes that while it is natural to confront the manipulations of deceivers and self-deceivers with some suspicion, the motivations and the consequences of such deceptions can be at least benign, and even beneficial, for those involved.

Having attempted to set aside popular and most often simplistic assumptions about self-deception, Rorty then moves to explore what are genuinely complex strategies for effecting deceptions: distraction from what we do not want to see; the use of vagueness to compel ourselves into serious action; the creative reconstruel of descriptions to throw good light on a bad situation; the appeal to positive generalization to lift us out of the tedium of wearing details; or contrarily, the selective abstraction and attention upon a few positive details as respite from an overwhelmingly onerous chore; and so on. There are also second-order policies which aid and abet self-deception by allowing us to rationalize our manipulations. Compartmentalization allows us to entertain conflicting projects. The assignment of high social value to certain customary ways of acting allows us to get away with questionable personal habits. The construction of general theories of human nature and identity provide a screen for certain activities that shield them from closer scrutiny. We ignore the contingency that qualifies all of our intentions and activities and are happy to focus our attention on the anticipated "normal" outcome. We invent personal dispositions which we do not initially have in the hope that such constructions will serve our goals. There are often real traces of the self-deceiving self when self-deception is employed as a strategy for self-transformation, and yet practically speak-

ing, it is usually easier to affirm the socially sanctioned self we are trying to perform than it would be to justify the self we would leave behind.

There is a global kind of self-deception that enables us to overcome the uncertainty and doubts which attend a thoughtful understanding of our life in the world: what Hume would regard as a natural inclination to believe and to trust beyond the evidence that we should. This kind of self-deception is the standard filler in the psychological operation of believing and imagining, and we couldn't get along without it.

Our personal and particular self-deceptions are beneficial to us in enabling us to get on in our relationships and occupations by making them more tolerable than they really are. At a social level, self-deception is so essential to maintaining communal harmony and equilibrium that it is impossible to conceive of a society which would not actively encourage it.

Rorty wants to move away from seeing the individual as either primary agent or beneficiary of self-deception or belief. Self-deception and belief sometimes occur at a superpersonic level in identifying with a group or community, and sometimes, at least if we accept the analyses of some kinds of psychological theorists, at the level of relatively independent subsystems that make up our complex psychological organization.

One very real danger in appreciating the collusive nature of self-deception and of subscripting agents and beneficiaries, deceivers and deceived, is the problem of assigning appropriate responsibility. For Rorty, this concern about self-deception is of a piece with many other, at least superficially, rational activities, which in fact themselves rely on shadier intellectual and psychological activities to function at all. And our only safeguard is Socratic inquiry—"the permanent possibility of asking critically evaluative questions." Given the social context and complicity that attends self-deception, another practical tactic we can employ in protecting ourselves against the more pernicious forms of self-deception is being careful about the company we keep. Finally, we must rely upon our other established psychological and intellectual habits—a distaste for hypocrisy, for example—as well as luck, to keep us from being victims of our selves.

For Robert C. Solomon, it is not clearly or obviously the case that deception is always wrong. And it is neither clear nor obvious that what is "true" is self-evidently so. Given the historical failure in coming to terms with the very contingencies of truth, Solomon wonders why over the centuries it has retained its value as a grounding moral virtue. Ironically, the most persistent deceivers and self-deceivers are those who take their noble vocation to be the unrelenting pursuit and defense of truth itself: the philosophers.

To begin his discussion of deception as a philosophical issue, Solomon chooses to distance himself from analytic discussions of self-deception—the

paradox of asserting contradictory propositions—as being a rather profitless foray into a complex psychological and cultural landscape. To the extent that the self is an embedded social self, differing significantly in different cultures, "self" deception goes well beyond deceptions directed *at* the self. In surveying the record of truth in the Western philosophical tradition, Solomon arrives at what many philosophers (perhaps themselves self-deceivers) would regard as a rather perverse conclusion: the actual practical value of truth has been grossly overrated, and deception is a much undervalued mainstay of social solidarity. Moving from the Western experience, where many good minds have recognized the often negative value of accepting the very notion of 'truth,' Solomon observes that there are indeed alternative cultural sites where social harmony is a value that far outweighs the importance of simple truth-telling.

Solomon rehearses the three most familiar arguments against lying: the deontological moral law that proscribes lying under any circumstances; the utilitarian position that calculates the relative good or bad consequences of lying; and the virtue-based argument that honesty is the expression of a cultivated character. What makes this third position more interesting is the possible assumption that lying and (self-) deception can have a corrupting affect on intersubjective and communal relationships. Solomon wants to redirect attention from individualistic concerns with principles, consequences, and character, to deception as a social phenomenon that possibly undermines particular interpersonal commitments. Virtue ethicists tend generally to assume a personal autonomy and persistency that marginize the ongoing negotiation of interpersonal relationships in favor of tricky situations. The problem with lying as Solomon sees it is that it can betray trust in concrete, personal relations. However, when we consider deception at the level of specific personal interaction instead of at some more rarified theoretical level, we realize immediately that deceptions of one kind or another are so important in maintaining working relationships that blanket condemnation of interpersonal dishonesty is morally naive.

In relocating both deception and self-deception as often mutually entailing aspects of a socially embedded self-consciousness that informs all of our interpersonal relationships, Solomon is anxious that the complexity of the pattern of feelings and the tenuousness of the lines that demarcate them be fully appreciated. As such, nuanced sympathy and understanding are often more appropriate responses to deception than raw blame. Taken a step further, the entire fabric of our cultures is spun from inherited rituals and myths. This cultural mythology is our resource for communal meaning, and to "debunk" it would impoverish our lives utterly.

12

Turning to a philosophical interrogation of the phenomenon of deception, Solomon recommends a holistic approach that focuses attention not only on the victim of the deception, but also on the probably self-deception of the person who would deceive. Again, deception is a radial phenomenon, the repercussions of which spread from a single act to incite public cynicism and mistrust. Each of the traditional approaches to the phenomenon of deception give us a partial understanding: the Kantian deontological analysis provides some insight into the nature of lying; the utilitarian analysis looks to its consequences; and the virtue-based approach focuses on the character of the deceiver. An adequate philosophical understanding cannot be constructed by combining partial perspectives, but must begin from an appreciation of the interpenetrating and interdependent nature of all aspects of the phenomenon. It is the complexity of deception and its implications that makes it difficult to sustain a lie, and that gives self-deception an important supporting role. It is this same complexity that makes deception and self-deception omnipresent colorations in the myths and metaphors by which we live. While deception can at times be morally reprehensible, there are just too many instances where it is a factor in the underdeterminacy and productive vagueness of our interpersonal transactions to allow for any globalizing characterization.

Finally, how "universalizable" is any culturally specific understanding of deception, and particularly, self-deception, when notions of 'human agency' differ so markedly in different cultural locations? While Solomon would allow for radically different social constructions in other cultures (as posited by Ames among others in this same volume), he would argue with some enthusiasm that the current caricature of the Western liberal individualist self is a most unfortunate and simplistic misconstruel of a tradition that left Descartes far behind in its development of social sensibilities. We "Westerners" are much more socially constructed in our practices than our theories (and often our adumbration of our theories) would allow.

Solomon concludes by reiterating his distinction between the *internalist* and the *externalist* models of self-deception, based upon internalist and externalist conceptions of 'self.' The former is fundamentally an epistemological problem—knowing and yet seeming not to know—that describes a relationship between a persona and a set of beliefs. The resolution of the paradox is to suggest various alternative models of the architecture of the self that would explain away this contradiction.

The externalist self by contrast is embedded in a fluid set of roles and relationships, and externalist self-deception is an always-social performance which attempts to manipulate such roles and relations to given effect. Solomon's point is that regardless of what many of our philosophers might

13

want us to believe, our own working models of self and self-deception are probably considerably closer to the contextually based externalist conception than they are to the more abstracted internalist conception. This being the case, our greatest and most insidious self-deception is to believe and to act upon our myth of the autonomous individual.

Kathleen Marie Higgins joins Allen Wood in his critique of Jean-Paul Sartre's several models of "bath faith" as self-deception. The interrogation of Sartre's position is important because he is so often referenced in discussion of self-deception. Higgins points out that Sartre's typical case of bad faith depends upon a severe subject/object dualism. To begin with, given that bad faith is largely a condition of self-awareness, third-person reports such as Sartre's have marginal reliability. Again, the dependence of Sartre's models on conflicting propositional content makes self-deception into something much simpler than it seems to be, leaving out as it does a whole range of propositionally unspecificable feelings and attitudes. Rather than contradictory propositions, we more likely have open-ended circumstances that are in a constant process of complex renegotiation and compromise. This is particularly true when so many of Sartre's examples depend upon human sexuality, a rather multivalent aspect of the human condition that few would claim to understand, even with respect to their own feelings. The case of wives denying inescapably factual sexual pleasure by distracting themselves in the act of intercourse with their unfaithful husbands is at best a contrived and unconvincing third-person (indeed male) attempt to anticipate the self-awareness of another person.

Higgins moves on from Sartre to reassess both the nature and the moral culpability of self-deception. Consistent with Amélie Rorty and Allen Wood, Higgins chooses to define self-deception as part of an integrative process where selective inattention saves us from circumstances we would rather not confront. To explore the positive content of consciousness experienced by the self-deceiver, a concern unnoticed in Sartre's reflections on self-deception, Higgins introduces "kitsch": not only bad art, but bad art that often generates self-deception. Higgins uses Karsten Harries and Milan Kundera to define the term *kitsch* as ironically charged objects which, irrespective of aesthetic merit, have the power to elicit deep, often socially generic, emotions in the observer. Kitsch is not an object, but the function of an object as a shared human image to establish intersubjective emotional relations.

Kitsch differs from Sartrean self-deception in that disclosure of a particular positive atmosphere is more focused than what is concealed. It is emotional stimulation that can serve many ends—nationalism, propaganda, and so on—by filtering out all that is incompatible with the overwhelming

14

and totalizing feeling. It would appear that most of the concern expressed in Sartre's "bad faith" was directed at his own specific situation in which political apathy was an acceptable alternative to political responsibility. Higgins suggests that even with this concern, the kitsch explanation is more compelling than Sartre's assumption that particular propositional content is being denied in bad faith. At least the sometimes enjoyment of kitsch cannot be condemned with the same moral fervor that Sartre directs against bad faith. Higgins, then, arrives at a conclusion not altogether unlike that of William LaFleur in this same volume. The hunt for hypocrisy, to the extent that it entails an unquestioned and global interpretive scheme, can easily become a far more dangerous undertaking than the occasional, often therapeutic lapse into self-deception.

Barbara Herrnstein Smith, in "Unloading the Self-Refutation Charge," sets the problem: it is a common ploy of defenders of philosophical orthodoxy to accuse "the deniers, rejecters, and abandoners" of such views, of incoherence, exposing in the course of the debate how the positions of these relativists, skeptics, perspectivists, constructivists, postmodernists, and so on are self-refuting. Further, given the social goods that philosophical orthodoxy sustains and the evils that it denies, it is also characteristic of our theoretical conservatists, even prior to the demonstration of self-refutation, to demonize the theoretical innovators as morally suspect and intellectually capricious. The charge of self-refutation, while clear enough to the audience of sympathizers to whom it is directed, is often empty and irrelevant to the positions of those so exposed, signaling nonengagement and even incommensurability. Herrnstein Smith's stated purpose is not to defend any specific unorthodox views against the orthodoxy, but rather to examine the often questionable ways in which the charge itself is deployed in an effort to make it just a bit less automatic.

The discussion begins from the classic accusation of self-refutation in the *Theaetetus* where the "relativist" Protagoras is taken to task, setting the model for the exposé of our more recent subversives. The structure of the self-refutation is simple, and the dubious paraphrasing and inferences are obvious to most classical scholars. Even so, these same scholars are willing and able to redeem it by a sympathetic appeal to the nonnegotiable vocabulary of logical argument that saves the condemnation of Protagorean relativism. An important point that Herrnstein Smith makes here (and in several other recent publications) is to challenge the egalitarian fallacy supposed by many of the critics of Protagoras' position, namely that Protagoras and all such relativists in rejecting the absolutist and objectivist claims to validity are asserting that any opinion is as good as any other. This is a fallacy because, in Herrnstein Smith's own words, "if someone rejects the notion

15

of validity in the classic sense, what follows is not that she thinks *all* theories (etc.) are *equally* valid but that she thinks *no* theory (etc.) is valid *in the classic sense*." Theories can be evaluated on standards such as "applicability, connectibility, stability," and so on that depend on locating specific conditions of perspective, interpretation, judgment, and so on rather than on objectivist standards. The "anything goes" accusation is simply a restatement of "any opinion is as good as any other" and is properly dislodged by the countercharge of question begging that leads to impasse.

The heart of the matter, then, is that many philosophers of the more orthodox turn believe and further claim that there are given constraints— "prior, autonomous, transcendentally, presupposed, and (properly) universal"—characteristic of our cognitive activities and their conceptual and discursive products. Herrnstein Smith would assert (and many of the perspectives taken from other traditions represented in this same volume would confirm) that what are being posited as "inescapable presuppositions" and "truth absolute" are the effects of participation in a specific conceptual tradition and its idiom. Bringing this kind of universalism into a comparison with the Chinese tradition, as one example, would expose it as a peculiarly and sometimes pernicious Western ethnocentrism.

The theater of instruction and the players that we discover in the *Theaetetus* are mirrored in the modern academy: the bright student enthralled by dubious doctrines is, after undergoing a thorough cleansing often through public trial and humiliation, led to the light and confirmed in the community of the orthodox.

In the remainder of her essay, Herrnstein Smith offers an explanation for what she takes to be the "intriguing phenomenon" of "the orthodox believer's conviction that he believes what he does because it is true, while skeptics and heretics believe what they do because there is something wrong with them." While this kind of "cognitive conservation" provides us with "intellectual stability, consistency, reliability, and predictability," when taken to an extreme it leads to the dangerous assumption that one's convictions are undeniable, and no adequate alternative formulations are possible or even coherent.

The alternative to the "cognitive conservatist," honed in rationalist philosophy to a fine and ambitious art, is the modesty of the postmodern skeptic who, allows that the conceptual systems of classical epistemology operate well enough for some, but not for all, and that alternative schemes can be advanced and defended on the basis of other, always contingent criteria. This skeptic is not embarrassed or disturbed by the contingency that attends her and everyone else's favorite theories. Unfortunately, as she would allow the classical epistemologist a place among competing doc-

16

trines, all that the traditionalist can hear is her self-refutation on the basis of his presuppositions.

All is not lost. This play has had an enormously long run—a persistence that suggests an inherent instability that prevents real deadlock and incommensurability.

Joel Kupperman begins by rehearsing several of the possible models of self-deception and some of the accounts available in recent philosophical literature (such as Stephen White's new turn on homuncular subsystems), acknowledging that many of these solutions depend upon the application of an over-determined, and hence less than nuanced, psychological vocabulary that brings their adequacy into question. An important aspect of "desire" that Kupperman underscores is the generally acknowledged extent to which context influences behavior. What compounds the "situational" aspect of desire is the indefiniteness of most people's sense of self. This fluidity both of self and of situation leads Kupperman in the direction of Mark Johnston's understanding of self-deception as "subintentional tropisms" rather than deceitfulness among those homuncular subsystems that are ultimately constitutive of a self. Moving down this path, the major question that emerges (a question posed in the *Tao Te Ching* passage which is discussed at the beginning of this chapter) is that given this fluidity, is there any knowledge which can be contrasted with self-deception? In the *Tao Te Ching* passage, self-deception only becomes a possibility when better and worse behaviors are acknowledged and persons as a consequence begin "representing" (and "misrepresenting") themselves among these choices.

A further complication in the analysis of self-deception is the distinction between a rehearsed inner self-image and one's more external presentations. This distinction often serves a positive function, for a self-deceptive and exaggerated inner presentation can be an encouragement to do better in one's external presentation. Another dimension of the fluidity of desire comes with the Sartrean recognition that "I am what I am" (good faith which can also be bad faith if it is resignation) which stands in some tension with "I am what I choose to be" (good faith).

But what are the limits on what one chooses to be? What are one's live options? Whatever they are, it seems in Sartrean terms they would be deliberate adjustments in one's decisive yet unchosen original project. Having said this, Kupperman would hold out (with Confucius) that morality is more than following the right rules. What then is moral goodness? First, it is more than doing the "normal" thing. It is (*pace* Confucius) a kind of self-knowledge of a processional, and thus provisional self who knows who he genuinely is, and yet who is open to further articulation through self-reflection.

17

In part 1 of his essay, A. S. Cua uses a careful and nuanced analysis of the *Ta-hsüeh* 大學 (the Great Learning), one of the canonical *Four Books* defining classical Confucian philosophy, to reconstruct a vocabulary both of personal cultivation and of self-deception. Cua avails himself both of classical texts and of medieval commentaries to flesh out the cluster of terms necessary to bring the Confucian project of self-realization into focus and to identify the specifically Confucian species of self-deception. In attempting to allow the text to speak for itself, Cua wants to find an internal perspective that foregrounds the cultural significance of the classical Confucian conception of 'human agency' without overwriting it with more familiar Western models.

In part 2, Cua explores self-deception as an interpretative diagnostic concept for one's failure to achieve sincerity in one's thought (*ch'eng yi* 誠意). In Hsün Tzu, self-deception is a failure to take the entire picture into account. It is a kind of obscuration (*pi* 蔽)—a kind of darkening of the heart and mind—brought on by preoccupation with one aspect of a situation without a balanced assessement of the consequences. Obscuration is invariably motivated by desire of one kind or another and can only be rectified by wise and informed deliberation through which one clarifies the heart and mind.

In attempting to articulate a Confucian conception of 'self,' Cua begins from Stephen Toulmin's observation that entifications of self generally emerge by hypostatizing "self" as it occurs in everyday reflexive idioms: for example, *self-reproach* and *oneself*. In examining the reflexive binomials that occur in the classical Confucian texts, Cua discovers that almost without exception they have a diagnostic and interpretive function in the process of self-examination (*tzu-hsing* 自省) that necessarily attends the project of self-cultivation (*tzu-hsiu* 自修).

Attempting to bring the Confucian conception of 'self' into sharper focus, Cua borrows Hsün Tzu's distinction between generic terms (*kung ming* 共名) and specific terms (*pieh ming* 別名) to distinguish a free-floating abstraction from a concrete, reflexive usage of self that emerges in the context of practical discourse. Through a critical examination of the interpretive commentaries of other contemporary scholars—Herbert Fingarette, Tu Wei-ming and Roger T. Ames—Cua affirms the relational and inspirational character of this polymorphous "self" while at the same time insisting upon the crucial role reflexive *tzu* 自-locutions play in moving from generic to specific discussions of self.

Robert Cummings Neville begins from a comparative reflection on self as "self-reference" in representative thinkers of the Western tradition (Paul, Augustine, Hegel, Kierkgaard), and the curious absence of this same

18

notion of objectified and superordinated 'self' in the Confucian tradition. The engine that has driven a Western philosophical interest in self is the observed tension between self-contradiction and self-consistency. Neville singles out the cultivation of normatively efficacious responsiveness to one's particular world as the signature of the Confucian self, where this responsiveness operates on many different levels of discourse through physical, cognitive, and emotional habituation. Excellence in the complex of sign-structured behaviors is authenticity in its broadest sense: a "sincerity" that precludes sustained contradictions.

Turning to self-deception, the West owns three dominant species: inner psychological self-deception directed by the unconscious (Freud), ideological self-deception which emerges when one allows one's authentic self to be overwhelmed by membership in and identity with a particular class or group (Marx), and the self-deception that arises from ignorance of the world in which we live combined with ignorance of our own limitations (tragic self-deception). Likewise, Confucian self-deception can be organized under three major headings. In rather sharp contradiction to the popular Western notion of 'unitary, autonomous self,' the dominant model of self-deception in the Confucian tradition is precisely the selfishness entailed in the identification of and service to an isolated self. A mechanism designed to thwart the rationalization of selfishness is the deep commitment to seriousness and self-criticism.

The second Confucian model of self-deception arises from the inadequate performance of formal, ritually structured habits, a concern that is signaled in the tradition by an emphasis on self-awareness, caution, and self-criticism. The problem stems from a semiotic dysfunction where one's social vocabulary is insufficient to effectively communicate the desired relationship—one thinks one is performing signs of friendship but in fact is miscommunicating.

A third kind of self-deception is a consequence of ritual expression (*li* 禮) outstripping a cultivated humanity (*jen* 仁). Under such circumstances, form (*wen* 文) overshadows substance (*chih* 質), a situation unacceptable to basic Confucian sensibilities. It is in this sense that *jen* is more fundamental than *li*—a touchstone whereby to evaluate all personal, social, and political structures. When one's heart is not invested in the relationship, no amount of formal correctness is sufficient to secure the bond, where increasing reliance on ritual behavior only leads to pompousness.

Roger T. Ames begins by suggesting that self-deception as it is generally understood is a decidedly Western philosophical problem that has its origins in the "one-many" structure of classical metaphysics. The assumption of a "unitary self" necessary to make the problem coherent is an analog

19

to the familiar "one-behind-the-many" chain-of-being model that grounds much of classical Western thinking, making it both metaphysical and dualistic. As such, this superordinated self has little relevance to Chinese philosophy. In fact, given that person is irreducibly social and multivalent in Confucian philosophy, the kind of duplicity that has been a central concern in this tradition is a decidedly social phenomenon.

Ames introduces a "focal-field" notion of 'self' to establish a contrast with the one-many model and follows the Platonic attempt to define a viable conception of 'person' by appealing to the analogy of a viable state. While Plato's is an abstract theoretical structure with ambitious claims about the universal nature of humanity, Ames insists that the Chinese equivalent, in the absence of metaphysical assumptions, must be a specific, concrete political regime which is historically and culturally contingent. Using the dynamics of the formation of the Han dynasty as a model for the way in which a person in Confucian society comes to be constituted, Ames suggests that the radial structure of the political and intellectual character of the dynasty is shaped by and reflected in the overlapping radial structures of those persons who collectively constitute it.

Ames then moves from the specific historical example of the formation of the Han dynasty to suggest that this radial or focal sense of order is pervasive in the tradition, identifiable in the way in which Chinese culture has come to organize its world, its history, its canonical literature, and knowledge itself. Perhaps the most obvious illustration of this radial order is the ritual constituted family, and by extension, community. Each "self" is an abstraction from a constantly shifting matrix of hierarchical roles and relationships which make life within family and community meaningful by providing a social syntax that establishes both one's place and one's status.

In this focus-field alternative to the one-many model, there is no superordinated, unitary self. Ames uses his engagement with Herbert Fingarette in the published literature to clarify the point that "self" can only be a reflexive impulse within a repertory of socially embedded experiences, desires, and beliefs that are articulated through multiple levels of communal discourse. Self is a specific, open, and provisional pattern of social discourse. As such, it is always positively and productively underdetermined in the sense that the constitutive patterns of each person are responsive to changing environments, being renegotiated in each and every particular circumstance. The alternative to a teleologically driven model of self is a notion of 'self-sufficiency': a self which achieves personal harmony by maximizing its possibilities in each situation.

Coming to the Confucian equivalent of self-deception, then, Ames insists that it is "selves"-deception in the sense that Confucian self-

20