

## PART ONE: IDEALS

Because moral laws have to be valid for every rational being as such, they are to be derived from the general concept of a rational being as such, and in this way explicate all morality, which requires anthropology for its application to human beings, at first independently as pure philosophy, that is, entirely as metaphysics ...[G, Ak.412]

---

In Volume I of this project, I offered various arguments against the Humean conception of the self. Many of these were designed to call into question the legitimacy and veracity of the belief-desire model of motivation as a comprehensive and universally applicable account of moral motivation. These arguments did not question the obvious fact that desires sometimes do cause action. Rather, they questioned the assumption that *only* desires *can* cause actions. In this brief introduction to Part I my aim is to sketch an idealized account, not only of how reason can cause action; but of how reason always does cause action, even in those cases where desire precipitates it. That is, my aim is to outline the alternative, genuinely comprehensive and universally applicable reason-based model of motivation I develop in detail in Part I, relative to which the belief-desire model is merely a special case.

My arguments in the following five chapters are independent of and prior to any particular substantive moral theory, whether metaethical or normative, in which a Kantian conception of the self might figure. They have the same role in Kantian metaethics that the discussion in Chapters II through V of Volume I have in Humean metaethics: They aim to articulate certain foundational presuppositions in action theory, decision theory, and philosophical psychology about the nature of rationality that an adequate Kantian metaethics must presuppose. I apply this model to an account of specifically moral motivation in Chapter VI. Later chapters in Part II build on these foundations by developing criteria of adequacy for normative moral theory – without, however, attempting to make the case for one normative theory in particular over others with which it shares certain essential features in common.

In order to defend in detail this alternative, reason-based model of motivation, however, I need first to limn an alternative view of how and where in the structure of the self reason might operate. The unified account of reason I offer in Chapters II and III, following, functions simultaneously as a model of motivation and a model of rationality. It is in this sense that I claimed in Chapter I the proposed Kantian conception of the self to be simpler, prettier, and explanatorily more powerful than the prevailing Humean paradigm. Since this single, unified account of transpersonal rationality is merely an elaboration of the weak, canonical conception of

theoretical reason as defined and governed by the norms of deductive and inductive logic, it is also more deeply entrenched in our thinking than the Humean model that it is claimed instrumentally to serve. That model puts this one at the disposal of the unlimited range of contingent and variable ends particular agents adopt. Whereas none of those ends are necessary or indispensable, the canons of theoretical reason that enable their realization are both. This conception is therefore metaphysically and conceptually primary even within the Humean conception.

Whereas Hume regarded the canons of theoretical reason as mere propositional objects of calculation and computation for maximizing the satisfaction of desire, Kant maintained that the principles of theoretical reason structure the self by supplying necessary conditions for its unity. Kant thought that these principles set certain minimal requirements of logical consistency and coherence that all conscious experience must meet; and therefore that unified subjects and objects of experience must meet as well. Kant contended that any possible experience that failed to meet these requirements would be “nothing but a blind play of representations, that is, less even than a dream.” (1C, A 112)

Consider the implications. In Volume I, Chapter II, I argued that the revisionist, tautological conception of a desire was not robust enough to do the needed explanatory work, and therefore was no proper desire at all. I also offered a representational analysis of desire according to which some intentional state of the agent is a desire if it includes certain sorts of conscious experience of its intentional object. A desire, on this analysis, is a certain kind of complex experience. If conformity to the minimal consistency requirements of theoretical reason is a necessary condition for integration into a unified self, and if no possible experience that fails these requirements can form part of a unified self, then in particular no possible desire that fails to conform to them can form part of a unified self. On Kant’s analysis, a desire that fails the minimal consistency requirements of theoretical reason is “nothing but a blind play of representations, that is, less even than a dream.” Hence no object of desire that fails these requirements can precipitate action, because no such desire can be experienced (I discuss behavior precipitated by unconscious desires in Part II, Chapters VII and VIII).

To have a desire and pursue its satisfaction in action both presuppose the existence of a unified subject whose desire and action they are. In order to have and act on a desire of any kind, then, fulfillment of the necessary conditions for unified subjecthood must be presupposed. If Kant is right in maintaining that minimal consistency requirements of theoretical reason are among these necessary conditions, then no desire that fails those requirements can motivate action because no such desire is one a unified agent can have. Motivationally effective desires as well as the final ends that are their ultimate intentional objects, then, are subordinate to the minimal consistency requirements of theoretical reason on Kant’s view. I argue in Chapter II, following, that theoretical reason thus provides necessary conditions both of action and therefore of its final ends.

However, Kant's conception of the self implies even more than this. Kant's conception also implies that reason itself can precipitate action, independently of desire – and hence provides sufficient conditions of action as well. For in order that the minimal consistency requirements of theoretical reason filter out anomalous motives, emotions and thoughts from conscious unified experience, they must function effectively as sentinels, as gatekeepers of coherence that evaluate such possible experiences for inclusion in or exclusion from conscious awareness. That is, they function as motivationally effective cognitive norms that select from the array of external and internal information and experience the *content* of both latent and occurrent thought, belief, emotion, desire, intention, and sensation, for minimal internal consistency with those which already form and constitute the structure of the self and character of the agent. Otherwise stated, unified agents are overridingly disposed to preserve their own internal rational coherence in cognitive acts of rational content-selection.

Then consider those instances in which such minimally consistent desire is absent, but occurrent thought is present, the content of which is minimally consistent relative to the agent's other experiences and dispositions, and so meets the criteria of theoretical reason. Here there need be no mystery as to what moves the agent to perform a particular action. As we saw in Volume I, Chapter VII in discussing Nagel, occurrent thoughts, beliefs, rememberings, recognizing, and so on are cognitive events with causal efficacy no less than are desires. We can distinguish among such events only on the basis of their intentional content. More specifically, I show in Chapter V below that we can distinguish the motivationally effective from the motivationally ineffective occurrent cognitive events only on the basis of their intentional content. And more specifically still, it is the intentional content of such an event that decides its degree of motivational efficacy relative to the agent's other experiences and dispositions. Most specifically of all, an agent can be motivated by the intentional content of an occurrent thought or belief to perform an action that expresses that thought or belief, whether a desire is present or not, provided that this content motivationally overrides the intentional content of other, competing beliefs and/or desires. This intentional content is rational if it satisfies the two minimal consistency conditions of theoretical reason which I develop in Chapter II, Section 4 below, and elaborate further in Chapter III, following. In Part II of this volume I describe some of the many ways in which we maintain the appearance of rational integrity even when actual rational integrity has been violated.

So, to apply this thesis to an example of specifically moral motivation for which a viable Kantian conception must provide an analysis, a whistleblower can be moved to publicize her company's unethical practices by an occurrent cognitive experience that, while minimally consistent with her other experiences and dispositions, nevertheless violates or threatens her convictions about fair labor practices. This experienced threat to her conception of fairness can motivate her to take steps to restore fairness by redressing unfairness, even though a desire to blow the whistle, whether self- or other-directed, is nowhere to be found – provided that the intentional content of this experience outcompetes in urgency the content of other intentional

states she may also experience (for example, fear, self-seeking, greed, etc.). Whether it does or not depends, not on whether or not she has a "pro-attitude" toward fairness, but rather on how deeply embedded in the structure of the self the concept of fairness is for her. If it is very deeply a part of her, she will be moved to defend herself against assaults on it. Such whistleblower behavior would be a paradigm example of transpersonally rational motivation. Chapter VI.8 limns a psychological apparatus for explaining in greater detail how this could happen, and Chapter IX.8 offers a justification for why it ought to happen. Hence the above ruminations sketch only in very general outline the argumentative strategy of this first Part of the volume. Clearly there is a great deal more to be said.