

Published in *Philosophical Forum XXIV*, 1-3 (Fall-Spring 1992-93), 188-232; reprinted in Robin May Schott, Ed. *Feminist Interpretations of Immanuel Kant*, (University Park: Pennsylvania State University Press, 1997), 21-73; and in John P. Pittman, Ed. *African-American Perspectives and Philosophical Traditions*, (New York: Routledge, 1997)

Adrian M. S. Piper

Xenophobia and Kantian Rationalism (1991)^{*1}

Contemporary Kantian ethics has given a wide berth to Kant's analyses of reason and the self in the *Critique of Pure Reason*.² Perhaps this can be ascribed to P. F. Strawson's influential fulminations against Kant's transcendental psychology in *The Bounds of Sense*.³ Strawson's view was an expression – one of many – of a post-war behaviorist sensibility, in which the best conceptual analysis of interior mental life was no analysis at all. In recent years this sensibility has become increasingly anachronistic, both in ethics and in philosophy of mind, and is in need of reappraisal on these grounds alone.

The neglect by contemporary Kantian ethicists of Kant's first *Critique* has been particularly unfortunate. It forecloses a deeper understanding of Kant's own ethical views, and robs us of valuable resources for addressing contemporary issues in metaethics and applied moral philosophy. It is virtually impossible to understand Kant's conception of the categorical imperative in isolation from his account of reason in the first *Critique's* Transcendental Dialectic; or his distinction between autonomy and heteronomy in isolation from his inchoate but suggestive formulation of the Two Standpoints Thesis in the Solution to the Third Antinomy; or his elaboration of that thesis itself in Chapter III of the *Groundwork of the Metaphysic of Morals*⁴ in isolation from the chapter on Noumena and Phenomena, the Refutation of Idealism, and the Fourth Paralogism in the A Edition of the *Critique*. Of course this is not to deny that these concepts can be put to excellent and fruitful use independently of ascertaining what Kant himself meant by them.

Moreover, the first *Critique* offers a developed conception of the self that provides a needed resource for defending Kantian ethics against Anti-Rationalist criticisms, such as that it is too abstract, alienating, altruistic, or detached from ordinary personal concerns to guide actual human behavior. The conception of the

*© APRA Foundation Berlin

self to be found in the first *Critique* is, to be sure, a thoroughly rationalistic one that no Anti-Rationalist would accept. Its virtue, however, is to demonstrate convincingly that in ordinary personal concerns, as well as in the guidance of human behavior, the scope and influence of rationality is inescapable.

Corresponding to these two considerations, the purpose of this discussion is twofold. First, I want to shed some light on Kant's concept of personhood as rational agency, by situating it in the context of the first *Critique's* conception of the self as defined by its rational dispositions. I hope to suggest that this concept of personhood cannot be simply grafted onto an essentially Humean conception of the self that is inherently inimical to it, as I believe Rawls, Gewirth, and others have tried to do.⁵ Instead I will try to show how deeply embedded this concept of personhood is in Kant's conception of the self as rationally unified consciousness.

Second, I want to deploy this embedded concept of personhood as the basis for an analysis of the phenomenon of xenophobia. I focus on this phenomenon for two reasons. First, it is of particular concern for African-Americans. As unwelcome intruders in white America we are the objects of xenophobia on a daily basis. This pervasive fact of our experience conditions all of our social relations, and may itself engender a reciprocal form of xenophobia in self-defense. It is therefore doubly in our interests to understand this phenomenon and the defects in rationality it manifests. Second, Kant's conception of the self affords potent resources for understanding xenophobia as a special case of a more general cognitive phenomenon, namely the disposition to resist the intrusion of anomalous data of any kind into a conceptual scheme whose internal rational coherence is necessary for preserving a unified and rationally integrated self.

I begin by limning the conception of the self as rationally unified consciousness I want to defend on Kant's behalf. This conception differs from Kant's actual pronouncements in only one respect: I incorporate Strawson's suggestion that, among the candidates for innate concepts in Kant's Tables in the *Metaphysical Deduction*, only the subject-predicate relation can be understood as what Kant would call a transcendental concept or judgment-form. On this view, all other such concepts are empirical, including that of causality. I then formulate the issue of the relation between transcendental and empirical concepts or categories and its relevance to an analysis of xenophobia. Kant claims that anomalous data that fail to conform to the transcendental concepts of the understanding cannot be experienced

by a unified self at all. Xenophobia is fear, not of strangers generally, but rather of a certain kind of stranger, namely those who do not conform to one's preconceptions about how persons ought to look or behave. It is therefore a paradigm case of resistance to the intrusion of anomalous data into an internally coherent conceptual scheme – a threat to the unity of the self defined by it. If a disposition to these preconceptions is innate, then xenophobia is a hard-wired, incorrigible reaction to a threat to the rational integrity of the self. If, on the other hand, a disposition to these preconceptions is the result of empirical conditioning, then xenophobia is corrigible in light of empirical data that may be realistically expected to compel the revision of those concepts.

In Section II I begin the exegetical part of this project by sifting through Kant's own claims about the relation between transcendental and empirical concepts, and conclude that empirical concepts, on Kant's view, instantiate transcendental ones.⁶ In Section III I locate Kant's concept of personhood relative to the distinction between transcendental and empirical concepts by arguing that this concept has both transcendent and transcendental status for Kant. This implies that Kant's concept of personhood is innate, and not subject to empirical revision. However, the way in which this concept is instantiated or applied is not similarly fixed. On Kant's account, we identify others as persons on the basis of our own self-identification as persons; and Kant insists that the only self to which we have epistemic access is empirical. In Section IV I examine Kant's account of self-knowledge and argue that Kant's distinction between noumenal and empirical selves does not foreclose veridical identification of oneself as a person. I conclude that Kant's transcendent concept of personhood is instantiated by particular empirical exemplars of personhood, i.e. particular persons with particular personalities, among whom each of us necessarily identifies oneself and only contingently identifies others.

In Section V I then deploy Kant's concept of personhood and his distinction between transcendental and empirical concepts in the service of a detailed analysis of xenophobia. I argue that it is a self-protective reaction to violation of one's empirical conception of people, and involves a cognitive failure to apply the transcendent concept of personhood consistently across all relevant cases. I try to show that racism, sexism, homophobia, anti-Semitism, etc. are pseudorational responses to xenophobia that depend on the mechanisms of rationalization, dissociation, and denial; and on a deep personal investment in the resulting

honorific stereotype of the valued group to which one belongs. Derogatory racial, gender or ethnic stereotyping of others, on this view, is a reciprocally interdependent consequence of honorific stereotyping of oneself.

Finally, in Section VI I recur to the text, in order to settle the question of the cognitive status of xenophobia within Kant's theory. I offer two interpretations of Kant's requirement that all data of experience conform to categories constitutive of the rationally unified self. Interpretation (A) demands that all such data conform both to transcendental and to empirical concepts, whereas interpretation (B) requires that they conform only to the transcendental ones. If (A) is correct, then another who is anomalous with respect to one's empirical conception of people cannot be a person for one at all. So xenophobia is incorrigible. But if (B) is correct, then another might violate one's empirical conception of people but be nevertheless recognizable as instantiating one's transcendent concept of personhood. So it would be possible to recognize the other as a person even though she violated one's empirical presuppositions about what and who people are. In that case, even if a disposition to xenophobia were innate, particular manifestations of it would be the result of conditioning and therefore susceptible to empirical modification. I examine the textual evidence for each interpretation, and conclude that (B) is correct; and that Kant's conception of reason as theory-construction implies resources within the structure of the self for overcoming xenophobia - resources frequently overshadowed, however, by empirical conditioning.

I. Kantian Rationalism

In the first *Critique*, Kant tells us repeatedly that if a perception does not conform to the fundamental categories of thought that ensure the unity and coherence of the self, they cannot be part of our experience at all. (A 112, 122, and B 132, 134)⁷ Kant describes these fundamental categories as "*a priori* transcendental concepts of understanding," by which he means innate rules of cognitive organization that any coherent, conscious experience must presuppose. The table of transcendental categories he offers in the *Metaphysical Deduction* are drawn largely from Aristotle, with considerable additional tinkering by Kant. They include substance, totality, reality, possibility, causality, and community, to name just a few. Some commentators have rightfully concluded that the most significant candidate for this elevated cognitive status is the subject-predicate relation in logic, from which

Kant derives the relational category of substance and property in the Table of Categories (Kant regards this as the result of fleshing out the subject-predicate relation or "judgment form" with "transcendental content," i.e. the sensory data our experience presupposes rather than the sensations we perceive as a result of it. (A 70/B 95-A 79/B 105)).⁸ The idea, then, would be that organizing sensory data in terms of this relation is a necessary condition of experience. On this view, if we do not experience something in a way that enables us to make sense of it by identifying properties of it – for example, in propositions such as,

That car is dark red,

or

I am tired,

we cannot consciously experience that thing at all.

This thesis – call it the *Kantian Rationalism Thesis* – has the merit of plausibility over the archaic list of categories Kant originally furnished. It does not seem too controversial to suppose that any viable system of concepts should enable its user to identify states of affairs by their properties, since concepts just are of corresponding properties, and to ascribe a property to an object just is to subsume that object under the corresponding concept. So any system of concepts should enable its user to ascribe to objects those properties of which she has concepts. The Kantian rationalism thesis – henceforth the KRT – is so weak that it may even be defensible in the face of anthropological evidence that languages considerably remote from Indo-European ones evince a cognitive structuring to the user's experience that is so different from our own as to be almost unintelligible to us. It would be an argument in favor of the KRT if it could be shown that the subject-predicate relation held regardless of the other ways in which culturally specific conceptual organizations of experience differed among themselves.

More precisely formulated, then, the KRT says that if we do not experience something in such a way as to allow us to make sense of it in terms of a set of coherent concepts that structure our experience, *whatever those concepts are*, we cannot consciously experience that thing at all. On this thesis the innate capacity would consist in a disposition to structure experience conceptually as such, but not necessarily to do so in accordance with any particular list of concepts,⁹ provided that the particular, culturally specific set *S* of concepts $c_1, c_2, c_3, \dots, c_n$ that did so satisfied the following requirements:

- (A) S observes the law of noncontradiction, i.e. the members of S are internally and mutually consistent in their application;
- (B) Any particular c_i in S is either
 - (1) an instantiation of some other c_j in S; or
 - (2) instantiated by some other c_k in S;
 i.e. S is minimally coherent;
- (C) For any cognitively available particular p, there is a c_k in S that p instantiates.

The suggestion would be that we can understand particular states of affairs only if (A) the concepts by which we recognize them are neither internally nor mutually contradictory; (B) those concepts are minimally coherent with one another in that each particular identified by them satisfies the subject-predicate relationship with respect to at least one other of them; and (C) that particular itself instantiates at least one of them. I develop this suggestion at length elsewhere.¹⁰ It says, roughly, that in order for something to register as a conscious experience at all for us, we have to be able to make sense of it in terms of some such concepts in the set; and that if we can't, it won't.

Suppose, for example, that we were to be confronted with some particular such that the concepts it instantiates satisfied (A) but violated (B), i.e. such that we could invoke a concept in identifying it consistently with the application of our other concepts; but that that concept itself bore no instantiation-relation to others in the set (i.e. aside from that of being a concept in the set). In this case, that which we invoked as a "concept" would in fact not be one at all, since the corresponding predicate would by definition denote only the single state of affairs it had been invoked to identify. Since there would be no further concepts in terms of which we might understand the meaning of that denoting term, it could not enter into any analytic truths. In short, this would be like cooking up a special noise to denote only one state of affairs on the single occasion of its occurrence. The enterprises of denotation and meaning themselves would fail.

Alternately, imagine what it would be like to be confronted by a particular such that its concept satisfied (B) but not (A), i.e. such that it enabled us to identify its properties in terms of concepts in the set, but the application of those concepts themselves was internally or mutually inconsistent. In that event, it would be possible to ascribe to the thing the conjunction of some predicate F and some other

one, G, that implied the negation of F.¹¹ Again the enterprise of identification itself would fail. If we were finally to fail to identify the thing or state of affairs in question as having a consistent set of properties, we would fail to identify it altogether. And then it could not be part of our conscious experience. If such cases characterized all of our encounters with the world, we would have no experiences of it at all and therefore no unified sense of self either.

These are the sorts of failures Kant has in mind when he avers, in the A Deduction, that

without such unity, which has its rule *a priori*, and which subjects appearances to it, thoroughgoing, universal, and therefore necessary unity of consciousness would not be found in the manifold of perceptions. These would then not belong to any experience, therefore would be without object, and nothing but a blind play of representations, that is, less even than a dream. (A 112)

Kant is saying that if we do not organize cognitively the data of our senses according to consistent and coherent rules, we cannot be rationally unified subjects. "For otherwise," he adds in the B Deduction, "I would have as many-colored and diverse a self as I have representations of which I am conscious." (B 134) I would, that is, lack a sense of myself as the subject in whose consciousness those representations occur. For a Kantian rationalist, then, the cognitive organization of experience according to consistent and coherent concepts is a necessary condition of being a rationally unified subject.

The KRT as explicated claims that only the subject-predicate relation counts as what Kant would call a transcendental concept or judgment-form; all the rest are empirical. Empirical concepts may differ as to how deeply entrenched in our cognitive dispositions they are. But all empirical concepts, for Kant, apply to and are formed in response to particular empirical contexts, rather than being necessary preconditions of experience itself. However, Kant did not devote sufficient attention to explaining the relation between empirical and transcendental concepts. If empirical concepts are contingent rather than necessary determinants of experience, then presumably we might have a particular experience even though we lacked one particular empirical concept by which to make sense of it – i.e. in the case where we had some other, nonequivalent empirical concept that did the job equally well. And Kant is silent on the question of whether we might have a particular experience that

conformed to the transcendental concepts but to none of our empirical concepts – for instance, of an empirical state of affairs for the evident properties of which we could find absolutely no fitting predicates at all. Is the formation of empirical concepts itself a necessary precondition of experience? Or is it as contingent on circumstance as those empirical concepts themselves are? Nor does Kant explain how susceptible to change our empirical concepts are, in light of their relation to transcendental ones on the one hand and to new or anomalous empirical data on the other.

These issues are central to the topic of this essay. Thomas Kuhn has documented the inherent impediments to paradigm shift in the natural sciences - their conservatism and constitutional insensitivity to the significance of new data, and their resistance to revising deeply entrenched theories in light of experimental anomaly.¹² Elsewhere I have argued that the resistance to integrating anomaly is a general feature of human intellection that attempts to satisfy a Kantian requirement of rational self-preservation.¹³ I have also offered elsewhere a Kantian analysis of a certain brand of xenophobic resistance to anomaly that finds typical expression in racism, sexism, anti-Semitism, class elitism, and homophobia, among other types of discrimination.¹⁴ The question at issue here is whether a Kantian conception of the self explains xenophobia as a necessary or a contingent attribute of the self; i.e. whether it is a hard-wired disposition to defend the self against attacks on its internal integrity that is impervious to modification, or whether a xenophobic fear of strangers as violating one's conceptual presuppositions about persons is contingent on such empirical conditions as upbringing, degree of exposure to diversity or integration, and peer-group reinforcement – and therefore revisable in light of new experience. Ultimately I think Kant's view implies the latter. This what I will try to show in what follows.

II. Transcendental and Empirical Concepts

Kant says many things about the relation between transcendental and empirical concepts, most of which are inconclusive. He says that empirical concepts are based on transcendental ones (A 111), that they are grounded in transcendental ones (A 113), that they are subject to them (B 163), that they must agree with them (B 164), and that their source is in them (A 127). None of this is precise enough to shed light on the actual relation between them. A more specific but fallacious account of the relationship is suggested by Kant's assertion that empirical concepts are a

consequence of transcendental ones (A 114). Regardless of whether by "consequence" Kant means "causal consequence" or "logical consequence," he clearly should not have said this: Causality as itself a transcendental concept is not a relation that can be maintained to hold between transcendental and empirical concepts, and transcendental concepts cannot be supposed to imply empirical concepts (indeed, Kant as much as acknowledges this when he says later that empirical laws cannot derive their origin from transcendental concepts (A 128, B 165)). Nor is the extended account of the reproductive imagination in the A Deduction at A 119-124 and in the B Deduction at B 152 helpful in ascertaining exactly in what the relationship consists.

Surprisingly, Kant does admit the existence of "derivative pure *a priori* concepts." These are derived by combining the transcendental ones with one another or with "modes of pure sensibility," i.e. our innate disposition to structure our experience spatiotemporally. Whatever the character of this latter process of combination, it cannot be identical to or even very much like that involved in schematizing the categories in time, since this, Kant tells, is what gives the transcendental concepts applicability to our sensory and spatiotemporal experience (A 140/B 179 – A 142/B 181); it does not engender any derivative ones. Among these derivative necessary concepts are those of action, passion, and force, derived from the transcendental concept of causality; and the concept of presence and resistance from that of community. (A 82/B 108) The concepts of action and passion are of particular interest for understanding the role of human agency and inclination in Kant's moral philosophy, and it is useful to see them identified as necessary preconditions of experience so early on in the first *Critique*.¹⁵ But Kant defers the project of enumerating all of these additional transcendental concepts to another occasion, and says nothing more about the nature of their process of derivation.

More helpful is Kant's assertion, in both the A and the B Editions of the first *Critique*, that empirical concepts depend on transcendental ones (A 114, B 164). This implies that transcendental concepts are a necessary but not sufficient condition of empirical ones, i.e. that transcendental ones make the empirical ones possible without ensuring any particular set of them. Earlier, in the *Transcendental Logic*, Kant has explained why transcendental concepts alone, i.e. in their corresponding logical forms of judgment, cannot furnish a sufficient condition of empirical truth. Although they do furnish criteria of logical truth,

[t]hese criteria ... concern only the form of truth, that is, of thought in general; and in so far they are quite correct, but not sufficient. For although our knowledge may be in complete accordance with logical form, that is, may not contradict itself, it is still possible that it may be in contradiction with the object. (A 59/B 84)

Here Kant observes that the fact that a proposition may satisfy logical requirements does not by itself determine its content; indeed, it may happen that a system of propositions may satisfy these requirements, yet its content might be "contradicted," i.e. conclusively disconfirmed by the objective states of affairs it purports to denote. (Also see B 190, A 155-157)

The sufficient condition for the veracity of empirical concepts - i.e. that which ensures the consistent and coherent application of at least one specifiable *kind* of empirical concept rather than any other to a particular ("rabbit" or "gavagai" rather than "H₂O" to small furry entities with long ears, for instance) – is given by the source of their transcendental content. That transcendental content itself is what Kant calls the manifold, and he thinks that the process of synthesizing or unifying the manifold under concepts is what specifies their content:

Synthesis of the manifold (whether empirical or *a priori*) is what first produces a cognition, which certainly may be crude and confused at first and therefore in need of analysis. But synthesis alone is what actually collects the elements into a cognition, and unifies them into a particular content. (B 103/A 78)

The crucial missing link in these remarks is an answer to the question of whether a randomly chosen element of manifold has attributes that lead us to collect it under one concept rather than another (i.e. whether natural kinds exist), or whether its attributes are conferred solely by the concepts that subsume it. Here Kant is silent on which account of transcendental content is correct, but later relies on the latter possibility to justify the need for a Schematism of the Pure Understanding. (B 177/A 138) This latter possibility implies that any datum could be subsumed under any concept arbitrarily, and therefore that there was no systematic relation between our capacities of cognitive organization and the particular data we organized – hence the need for time as a schema that mediates between them. By contrast, the former implies that these data carry markers or clues to the concept that most appropriately subsumes it, and therefore that there is at least some minimal correspondence between the way we organize the world and the elements of the world that we

organize. Kant furnishes no unambiguous evidence for this possibility (but see Footnote 17, below).

And what about the source of this transcendental content? Its nature and ontological plausibility is a point of endless debate among Kant scholars. Here I will simply state (but not defend) the view that this source is what Kant describes, in a passage in the Schematism willfully mistranslated by Kemp-Smith, as "the transcendental matter of all objects as things in themselves." (A 143/B 182¹⁶; also see A 20, A 28, A 30/B 46, B 60, B 75, B 125, B 145, A 168, A 223, A 372, A 375, A 385, B 422a-423, A 581).¹⁷ This is to suggest that first, in addition to our innate capacities for cognitive organization, a multiplicity of objects that are ontologically independent of us must provide us with the sensory data we organize, in order for empirical experience to occur; and second, if these objects provide such data, and our cognitive capacities are in good working order, systematically related empirical experience will occur.

This suggestion does not imply that the systematic relation we detect among empirical objects is identical to any relation that might obtain among ontologically independent objects that are by definition inaccessible to our empirical experience. Nor does it imply that the relation between these two sets of relation is one of causality as we now understand that term (although of course it might be). But it does imply that we are justified in thinking of the systematic coherence we discover in the empirical world we experience as at least a clue to the character of the coherence that may be presumed to actually exist among ontologically independent states of affairs as they are in fact. Indeed, in *The Regulative Employment of the Ideas of Pure Reason* Kant himself concedes this when, in arguing that the unity in nature we discover by exercising our rational capacities in inquiry and research is a necessary precondition of experience rather than a contingent outcome of it, he says,

In fact it is hard to see how there can be a logical principle of the rational unity of rules, unless a transcendental principle is also presupposed, through which such a *systematic unity is a priori assumed to be necessarily attached to the objects themselves*. ... The law of reason which requires us to seek this unity is a necessary law, since without it we would have no reason at all, and without this no coherent use of the understanding, and in the absence of this no *sufficient criterion of empirical truth*. So in reference to this criterion we must

necessarily presuppose the systematic unity of nature as objectively valid and necessary throughout. (A 650/B 678-A 651/B 679; italics added).

Kant's point is that we are required, by the fact that reason and understanding must unify all of our experience under increasingly inclusive concepts in order for us to have experience at all, to conceive of all the possible objects of experience thus unified – i.e. the empirical world of nature *in toto* – as an ontologically independent system that is necessarily unified as well. Kant is not claiming that we detect that objective system in our necessarily limited experience of the natural world as it appears to us. He is not even claiming that we can infer any veridical characteristics of it from the areas of systematic structure we do empirically detect. He is claiming only that we must rationally conceive the totality of the natural world as an ontologically independent, systematically unified whole in order to experience any empirical part of it coherently; and the passages cited earlier demonstrate his own thinking as an example of this requirement.

The sufficient condition for the veracity of empirical concepts, then – that which prevents conclusive disconfirmation (or "contradiction") of our explanatory theories by the objects we experience – is the natural world conceived as an ontologically independent phenomenon, unified under a maximally inclusive explanatory theory that can account for them, which we must presuppose in order to experience any of its natural objects at all. This assumption explains why, although an explanatory theory of the natural world might undergo revision in light of new empirical data, it can never be conclusively disconfirmed or "contradicted" by that data all at once. The more inclusive and sophisticated the theory becomes, the more anomalous data offer the challenge of revising and extending that theory into an even more powerful one that can integrate them, and the less susceptible the theory becomes to conclusive disproof by piecemeal anomalous evidence.¹⁸ Only relatively primitive or provincial explanatory theories are vulnerable to the kind of attack Kant describes. I consider some in greater detail in Section V., below, and argue elsewhere that the dogmatism with which such a theory is maintained is an index of its explanatory fragility.¹⁹ The external natural world, then, conceived as an ontologically independent, systematically unified and fully explicable whole, supplies the sufficient condition of empirical truth for Kant.

But now conjoin this suggestion to Kant's further claim that any conscious experience we have must conform to the transcendental concepts (B 162), and that

empirical concepts are "special determinants" of transcendental ones (A 126, A 128). From this latter claim we can infer that an experience that conforms to empirical concepts thereby conforms in content to the transcendental ones that determine them. Now Kant does not say whether by the word "determine" he means "designate" or "ascertain" or "specify", nor does the German (*bestimmen*) enlighten us on this question.²⁰ I will assume that by "determine," when used in this context, Kant means "specify" since it is a broader term that can be used more or less synonymously with either "designate" or "ascertain" in most contexts. Kant is then saying, in the above-cited passages, that empirical concepts specify more precisely some of the same content that is structured by transcendental ones and initially generated by systematically related things in themselves.

Again this does not give us direct access to the nature of things in themselves, since the content of a more inclusive concept can be specified in a variety of nonequivalent ways by less inclusive ones. But the specification relation between transcendental and empirical concepts preserves relevant content from systematically interrelated things in themselves through transcendental and then empirical concepts that give it cognitive structure. And we can think of the specification relation as for present purposes equivalent to the instantiation relation described in Section I.B.1, such that

(A) c_i specifies c_j if and only if c_i instantiates c_j .

On this reading, empirical concepts instantiate transcendental ones, and the more inclusive properties corresponding to transcendental ones may be ascribed to the less inclusive properties corresponding to empirical ones – as, for example, the property of being a cause of change may be ascribed to a behavior, which in turn may be ascribed to an action. This reading accords with the KRT, according to which each (empirically contingent) concept within a subject's experience instantiates the (transcendentally necessary) subject-predicate relation or judgment-form relative to some others.

III. The Concept of Personhood

The question of whether a Kantian conception of the self contains the resources for explaining and reforming xenophobia now can be reformulated more precisely as the question of whether an instantiation interpretation of the relation between concepts necessary for experience and those contingent to it can explain the

phenomenon of xenophobia with respect to its degree of conceptual entrenchedness and corresponding amenability to rational correction. Since xenophobia involves withholding recognition of personhood from those perceived as empirically different or anomalous, part of the answer to our question will turn on ascertaining the cognitive status of the concept of personhood in Kant's epistemology. To do this we first need to understand Kant's conception of the relation between transcendental and transcendent concepts.

Kant regards the "transcendental judgment-forms" as having two separate functions. The first, already discussed, is to structure cognitively the sensory data that unified experience presuppose. In that role Kant calls them "transcendental concepts" or "categories." But these logical forms – which under the KRT reduce to the subject-predicate relation – have a second function as well. This is to reason, construct syllogisms and hypotheses, formulate theories, and make deductive inferences at increasingly abstract and inclusive conceptual levels from the unified experience thus structured. In this way these judgment-forms not only unify experience according to certain innate cognitive patterns, but unify the resulting multiplicity of unified experiences themselves under more abstract concepts and theories according to the same basic cognitive patterns:

Understanding may be considered a faculty of the unity of appearances by means of rules, and similarly reason is the faculty of the unity of the rules of understanding under principles. Accordingly, reason never applies directly to experience or to any object, but to understanding, in order to give to its manifold cognitions a unity *a priori* through concepts, a unity which may be called the unity of reason, and which is of a very different kind from that which can be accomplished by the understanding. (A 302/B 359) ... In fact multiplicity of rules and unity of principles is a requirement of reason, in order to bring the understanding into thoroughgoing coherence with itself ... But such a principle ... is merely a subjective law for the management of the resources of our understanding, in order to reduce their general use to the smallest possible number through comparison of its concepts ... (A 305/B 362 – A 306/B 363)

In this second function, Kant refers to the transcendental judgment-forms as "*transcendent* concepts" or "ideas" of reason. Whereas the term "transcendental" refers to the necessary preconditions of experience, "transcendent" refers to that

which exceeds or surpasses the limits of experience. Kant's notion is that abstract theories (whether moral, psychological, theological, or cosmological) that unify all the relevant data under a minimum of explanatory principles necessarily transcend in scope of application the contingent and piecemeal data that empirically confirm them. Thus the difference between transcendental concepts of the understanding and transcendent concepts or ideas of reason is ultimately a difference in degree of abstraction from experience rather than a difference in kind. (Compare Kant's account of judgment as knowledge by subsumption at A 68/B 93 – A 70/B 95 with his account of knowledge from principles at A 300/B 357; also see A viii, A 302/B 359, A 311/B 368, A 329/B 386, A 409, A 643/B 671-A 644/B 672, A 651/B 679)

As a transcendent concept (or idea), the concept of personhood gives coherence to our occasional, particular empirical experiences of these characteristics of human behavior, by unifying them under this more abstract and inclusive notion that surpasses in scope of application any particular instance of human behavior that conforms to it. It thereby contributes to a standing expectation that other human beings will regularly behave as persons no matter how frequently this expectation is violated in fact.²¹ As a transcendental concept, by contrast, it is what makes our particular empirical experiences of these same characteristics of human behavior possible. It is what enables us to recognize particular occurrences of consciousness, thought, rationality, and action for what they are. Whereas the transcendent concept of personhood supplies us with a higher-level *conception* of what being a person involves – a standing conception to which particular individuals may or may not conform on any given occasion, the transcendental concept of personhood enables us to recognize those occasions on which they do.

Kant clearly regards transcendent ideas of reason, like the transcendental categories of the understanding, as innate in the sense that reasoning beings are inevitably led to them by virtue of the categories of reasoning they use. But according to the KRT, only the subject-predicate relation, and so the substance-property transcendental category that corresponds to it, is a necessary condition of experience. From this it follows that only that transcendent idea of reason which is generated by the subject-predicate relation is similarly rationally inevitable, if any of them are.

Kant thinks the subject-predicate relation engenders the transcendent idea of a rationally unified, temporally continuous self as the content of the concept of

personhood. His explanation in the Paralogisms of why he thinks this is not the most convincing argument available. What he could have said is just that the "I think" accompanies all our other concepts and therefore is instantiated in them, whereas it itself instantiates only the yet more inclusive substance-property concept. Being rationally inevitable, it is either just as innate conceptually as is that relation for Kant, or else is at least very deeply entrenched. This would be to suggest something like a hard-wired disposition to recognize others of our own (human) kind.

Because unified consciousness and thought presuppose cognitive structuring by rational categories, and because Kant believes reason can be motivationally effective, the concept of *personhood* also may be supposed to include, in addition to rationally coherent and persisting consciousness, the capacity for action. This departs from Kant's usage in the first *Critique* somewhat, where he uses that term in discussing only the former properties. But in the *Groundwork* his characterization of a person as a rational being (Ac. 428) makes explicit the first *Critique's* connection between consciousness, rationality, and agency. Now there is no obstacle to conceiving of a being that has these properties but does not know that she does; indeed, Kant claims that this is precisely the human predicament, since genuine experiential knowledge of this topic is foreclosed to us. In order to prove the inevitability and so the transcendental necessity of this concept for human experience, Kant must show that we are disposed to identify ourselves as persons on the basis of evidence from which we cannot help but infer that we are.

Kant has plenty to say about each of the properties of personhood, and there is plenty of textual support for assigning the concept of personhood a transcendental as well as a transcendent status. For example, in the opening sections of the Transcendental Dialectic Kant treats the concept of virtue, which for him is part of the concept of a perfectly rational being (A 315/B 372, A 569/B 597) as a necessary practical idea of reason (A 317/B 374 – A 319/B 376) – i.e. one that can motivate action. This implies a corresponding concept with transcendental status. This implication is strengthened by Kant's explicit assertion, in his discussion of the Platonic forms, that no human being coincides with the idea of what is most perfect in its kind, but nevertheless carries it – i.e. the idea of humanity – "in his soul as the archetype of his actions." (A 318/B 375) In *The Ideal of Pure Reason* Kant explicitly describes an "idea of perfect humanity" at A 568/B 596.

Similarly, Kant's treatment of the unity of the thinking subject as a transcendent concept or idea of reason that engenders the Paralogisms of Pure Reason implies transcendental status for the corresponding concept of the understanding, for example in the second paragraph at A 365. Indeed, he explicitly assigns that concept to the list of transcendental ones at A 341 / B 399, and says of it,

One quickly sees that this is the vehicle of absolutely all concepts, and therefore also of transcendental ones, and so is always conceived along with all of these, and therefore is itself equally transcendental.

A concept that is "always conceived along with" other concepts is instantiated by those other concepts. Therefore, empirical concepts instantiate more inclusive transcendent ones, and those, in turn, instantiate the most inclusive, highest-order transcendent concept of the thinking subject; and all of these instantiate the most inclusive transcendental concept of the substance-property or subject-predicate relation. This is just another way of suggesting – as Kant repeatedly does in the A and B Deductions – that the thinking self or "I" (as predicate) must be able to accompany (i.e. must be ascribable to) any experience. (A 116-117a, A 123, B 131-136, B 140; note, however, that he never explicitly acknowledges that the "I" would have to denote a property, not a substance) All of these passages taken together provide especially compelling support for the KRT.

Moreover, in the A Paralogisms Kant denies that the concept of a unified thinking subject can come from any empirical source when he declares,

Now I cannot have the slightest representation of a thinking being through any outer experience, but only through self-consciousness. So these sorts of objects are nothing more than the transference of this consciousness of mine to other things, which can be represented as thinking beings only in this way. (A 347/B 406)

He reiterates this at A 357, in the Second Paralogism. So he thinks that the concept of a unified thinking being is one I derive from the property of first-personal self-consciousness – a requisite, remember, for unified selfhood – and then ascribe to certain external empirical objects.

But which ones? How do I manage to ascertain which, among the array of available empirical objects, actually has that property, since I can find no empirical representation of it? Kant provides an answer to this question in the Solution to the Third Antinomy. There he first develops at length the thesis that the behavior of

actual human beings is subject to empirical laws of causality in the natural world, such that our behavior from a third-personal, observational perspective is entirely predictable: "[I]f we could exhaustively investigate all the appearances of the human power of choice," he tells us, "there would not be found a single human action which we could not predict with certainty, ... if, that is to say, we are merely *observing*, and, as happens in anthropology, want to investigate physiologically the motive causes of someone's actions." (A 550/B 578)

But he also says that the situation is different from a first-personal perspective:

Only a human being, who knows all of nature otherwise solely through the senses, also knows himself through pure apperception, and, indeed, in acts and inner determinations which he cannot class with impressions of the senses (A 546/B 574) [W]hen we consider the same actions in relation to reason ... in so far as reason is the cause *producing* them ..., we find a rule and order altogether different from the order of nature (A 550/B578),

i.e. the rule and order of rationality; of reasoning and deliberating about actions and goals, of forming generalizations and making inferences, and of reaching conclusions about what is the case and what to do that move us to act accordingly. From the first-personal perspective, then, we are not just objects of empirical investigation, determined by causal forces, but thinking and reasoning persons, determining our own actions through rational intention and will.

Thus the rule and order of rationality "shows, in its effects in appearance, a rule in accordance with which we may surmise rational motives and the kind and degrees of actions themselves, and judge subjective principles of the power of choice." (A 549/B 578) So Kant's solution to the problem of other minds, i.e. of how we can distinguish those third-personally observed objects who are similarly thinking subjects from those which are not, is to point out that although the behavior of all such objects satisfies the laws of causality, only that of some also may satisfy the laws of rationality. Only some external objects, that is, exhibit the capacity for rational *action*. The concept of a unified thinking being is, then, a transcendent and a transcendental concept we apply to those external empirical objects whose behavior gives evidence of being governed by the same laws of rationality we first-personally experience as governing our own.

IV. Self-Knowledge

Now Kant does not think such "acts" of first-personal introspection give us knowledge of ourselves as we are in ourselves, i.e. as noumena, but only as we appear to introspection. Moreover, Kant warns us about the dangers of transcendental illusion inherent in the use of reason when he says,

in our reason (considered subjectively as a human faculty of knowledge) there are basic rules and maxims of its use, which have all the look of objective principles, and through which it happens that the subjective necessity of a particular connection of our concepts is, to the advantage of the understanding, taken for an objective necessity in the determination of things in themselves. (A 297/B 353)

Kant is cautioning us against trusting our own, inescapable inference to the objective validity of our rational principles from their seeming necessity and universality. Our innate susceptibility to both of these conceptual traps raise the questions of whether our ascription of personhood to ourselves and others on the basis of our first-personal experience of the rule and order of rationality can have anything more than a contingent empirical foundation; and whether the concept of personhood itself can therefore claim any more validity than that.

Moreover, even if this concept should turn out to be necessary, it does not follow that we are necessarily justified in applying it to ourselves: if all we can know of ourselves is the way we appear to ourselves rather than the noumenal selves we are in actual fact, then the first-personal appearance of personhood may, for all we know, lack any basis in actual fact. Kant's distinction in Paragraph 25 of the B Deduction between what we can know – namely that of which we have empirical experience - and what we can consciously think or conceive – namely that about which the categories of thought enable us to reason, regardless of the extent to which it can be confirmed by experience – provides an answer to these questions.

Kant argues that it is the act of introspection (or – Kant's term – self-intuition (B 68-69, 153-156, 157-158a)) that enables one's self to appear to one at all. In the Transcendental Aesthetic Kant describes time as "the way in which the mind is affected through its own activity, namely by this situating (*setzen*) of its representation [in temporal relations], and so is affected by itself." This, he says, identifies time as an inner sense. (B 68) Kant's idea here is that the mind's cognitive process of forming and organizing representations itself causes the mind to situate

those representations in temporal succession; he elaborates on this idea at greater length in the Transcendental Deduction. A temporal succession of mental representations, however, is a property of the subject's interior consciousness, not of the external world; this is why Kant calls time an inner rather than an outer sense.

In the same passage, Kant goes on to say that the subject is itself the object of inner sense. What he means is that the temporal succession of mental representations that is a property of the subject's interior consciousness is identified by the subject as the subject's self when the subject turns its attention to it. (cf. also B 140) This succession of mental representations is the appearance of the self that Hume found when he looked within and searched in vain for the enduring soul or substance that had been supposed to unify these representations. Kant calls it the empirical self, in contradistinction to (A) the underlying noumenal self that does the appearing; and (B) the transcendental subject that both undergoes those cognitive processes and also has such properties (we will suppose (A) and (B) to be materially equivalent for purposes of this discussion). The empirical self is the self as it appears to one when one looks for it. It is therefore the product rather than the presupposition of these cognitive processes (cf. also B 152-3, 155 -6, 407). We can think of Kant's empirical self as equivalent to what I have elsewhere called one's *self-conception*.²² Kant describes this as an appearance of the self because, he claims, the very act of looking for it is what causes it to appear:

If the faculty of becoming self-conscious is to seek out (apprehend) what lies in the mind, it must affect the mind, and only in this way can it engender an intuition of itself ... it then intuits itself, not as it would immediately and self-actively represent itself, but rather in accordance with the way in which it is internally affected, and so as it appears to itself, not as it is. (B 69)

However, Kant later warns us at least twice about investing too much credence in our empirical selves, or self-conceptions, as a source of self-knowledge. The first time is in the Solution to the Third Antinomy. There he maintains that "[t]he real morality of actions (merit and guilt), even that of our own conduct, ... remains entirely hidden from us. Our imputations can be referred only to the empirical character." (A 551/B 579a) This warning is echoed in Chapter II of the *Groundwork of the Metaphysic of Morals* in Kant's remarks on the perniciousness and ubiquity of the "dear self." There he cautions us that

it often happens that in the keenest self-examination we find absolutely nothing except basic moral duty that could have been powerful enough to move us to this or that good action and so to greater sacrifice. But it cannot be ruled out with certainty that in fact no secret impulse of self-love, under the mere pretense of this idea, has been the actual, determining cause of the will. For this we gladly flatter ourselves, by falsely appropriating a nobler motivational basis. But in fact even the most strenuous probing of our hidden motives yields absolutely nothing, because when the issue is moral worth, it is not about the actions one sees, but rather about their internal principles one does not see. (Ac. 407; also see 419)

These caveats follow from Kant's previous remarks on the contingency and epistemic unreliability of the empirical self as a source of information about the transcendental subject to whom the empirical self appears. Here Kant is simply extending his remarks to cover the case of specifically moral self-knowledge as well. Thus the impossibility of knowing the noumenal self through acts of introspection would seem to foreclose reliance on first-personal consciousness, thought, rationality and action as conclusive evidence of authentic personhood in both first- and third-personal cases.

Kant also distinguishes sharply between the active spontaneity of the act of introspection, and the empirical self that is caused by this act to appear. He thus rules out direct and unmediated knowledge of oneself as an active and spontaneous intellect. *A fortiori*, he rules out direct experience of oneself as initiating the processes of reasoning and cogitation that would conclusively identify one as a person (the terms "experience" and "knowledge" for Kant are usually synonymous). Indeed, Kant's description in the *Groundwork* of the imperfect human will as one which, on the one hand, "is determined by reason," but on the other, "is not necessarily obedient to it by nature" or "subjective condition" (Ac. 413), suggests that reasoning and intellection are processes we experience ourselves as passively and sometimes resistantly undergoing rather than as actively initiating.

But Kant does acknowledge the possibility that a subject may nevertheless *represent* or conceive herself as an active, spontaneously reasoning and thinking subject (recall that this was the alternative he discarded in his account of self-intuition at B 69). The basis for this self-conception would not be the direct experience of active and spontaneous intellection. Instead it would be the

apparently spontaneous, uncompelled character of the content of those mental "acts and inner determinations" themselves. As Kant puts it,

I cannot determine my existence as a self-active being, but rather I represent to myself only the spontaneity of my thought, that is, of the determining, and my existence always remains sensibly determinable, that is as the existence of an appearance. But this spontaneity is why I call myself an intelligence. (B 158a)

Kant is, on the one hand, denying that I can ascertain my self-activity as a fact; but acknowledging, on the other, that I can conceive my thought as spontaneous in virtue of the autonomous character of my attempts cognitively to ascertain or specify things: it seems to me, that is, that my disposition to conceive and analyze are themselves self-initiated rather than externally caused or compelled. Although I can ascertain my existence only through empirical means, and therefore as an appearance, it is because I try to ascertain or determine things cognitively at all that I identify myself as an intelligent being.

So although the subject cannot know herself as an active intelligence, she can still represent herself as one, on the evidence of the autonomous quality of her thought. And she can include this representation among those constitutive of her empirical self-conception. Thus a subject may at least reliably conceive herself as having the properties that identify her as a person, even if she cannot experience herself as having them.

But Kant's distinction between what can be known and what can be thought or conceived is not only useful in formulating our self-conception as persons. It is also useful in formulating the evidence on which that self-conception depends. "The use of reason," he tells us, "is not always directed to the determination of the object, therefore to knowledge, but also to the determination of the subject and of its volition..." (B 166a) Following the suggestion in Section II that we translate *bestimmen* by the broad term "specify," Kant would then be saying that, in addition to ascertaining the nature of objects of knowledge, reason also can be used to shape the subject and its volition in certain specific ways. That is, reason can fix the form and specific content of the intentional object of the subject's will. It can fix the form of that intentional object in that it conceives the action that is its content as a valid conclusion of deductive and inductive reasoning. And reason can fix the content of that intentional object in that this reasoning process identifies a particular course of

action as the rational one to pursue. This interpretation conforms to the instantiation interpretation of the relation between transcendental and empirical concepts offered in Section II. So although reason alone may not yield knowledge of the true nature of the self, it may yield a precise and recognizably rational formulation of the subject's particular deliberations, resolutions, and intentions. Although this would not count as self-knowledge in Kant's technical sense, it certainly would constitute evidence for one's self-conception as a reasoning subject.

Further support for this reading can be gleaned from Kant's characterization of reason in the Solution to the Third Antinomy as atemporal and unaffected by empirical states, but as itself nevertheless a determining influence on them. (A 556/B 584) First, he describes the "appearances" of reason as "the ways in which it manifests itself in its effects." Presumably these effects are the particular, spatiotemporally specific instances of reasoning in accordance with the non-empirical canons of rationality. That Kant does not mean to identify the "effects" of pure reason with empirical action itself is clear from his earlier assertion in the same section that "the action, so far as it is to be ascribed to a way of thinking as its cause, does not thereby follow from it in accordance with empirical laws, that is, so that the conditions of pure reason precede it, but rather only so that their effects in the appearance of inner sense do." (A 551/B 579) If action is preceded by the effects of the conditions of pure reason in inner sense, it obviously cannot be identical with those effects. The only effects of pure reason in inner sense that can plausibly precede action are particular processes or occurrences of reasoning about what action to take. These empirical instances of valid reasoning specify the form and content of the intentional object of the will, as well as contribute to the motivational force of that will itself.

Second, at A 556/B 584 Kant also says of reason that it "is present in all human actions in all temporal circumstances and is always one and the same, but is not itself in time, nor falls somehow into a new state in which it was not before; it is determining, but not determinable in regard to this."²³ Here the idea is that the abstract canons of theoretical rationality themselves are not spatiotemporally local to any particular empirical action or situation, but are nevertheless locally instantiated by reasoning subjects who apply them to each such situation in such a way as to affect the action taken. The effect of reason on action cannot be merely to nudge it into existence causally as an occurrence. Instead it must affect action by specifying

or fixing the form and intentional content of the action relative to universal and necessary rationality requirements that the subject applies to all actions.

Third, Kant maintains that "when we say that in spite of his whole previous course of life the liar could have refrained, this means only that the lie is directly under the power of reason, and reason in its causality is not subordinated to any conditions of appearance or the passage of time." This means that an agent who is assumed to have been capable of doing otherwise is supposed not to have been handicapped in doing otherwise by intervening causal variables that might have obstructed the effect of reason on her action. This does not conflict with Kant's earlier claim that reason manifests itself in empirical appearances, if we understand by this that the abstract canons of reason are instantiated in particular, empirical occurrences of reasoning. Rather, it merely denies that there are any other internal empirical processes – such as inclinations or emotions – that might interfere with the subject's ability to recognize what reason abstractly requires of a specific instance of reasoning, or obstruct the effect of that specific reasoning process on action.

Finally, Kant concludes this paragraph by asserting that "although difference of time can indeed make a big difference in the relations among appearances, ... it can make no difference to the action in relation to reason." By this Kant means that irrespective of when the action occurs, and when the particular reasoning process that ought to precede it occurs, the action itself is nevertheless subject to evaluation in terms of rational criteria. This is something that the agent can recognize in so far as there are no cognitive obstructions to the "direct power of reason" to fix the form and content of her particular reasoning process.

Taken together, these passages contribute to an explanation of why conscious subjects who think, reason and act are inevitably led to identify themselves as *bona fide* reasoning and acting persons, even though they can have no knowledge of themselves as such. Centrally required for such identification is the subject's conception of reason as independent of and instantiated by particular empirical occurrences of reasoning that aspire to conform to it. This requirement is satisfied by the passages just considered. These provide evidence for adding to the conjunction of mental representations that constitute the empirical self the following:

- (1) the representation of the form and intentional content of one's deliberations and intentions as fully specified by abstract canons of theoretical reason (from B 166a);

- (2) the representation of particular empirical occurrences of reasoning as instantiations of these abstract canons of reason (from A 556/B 584);
- (3) the representation of these abstract canons of rationality as thereby causally affecting subsequent action (from A 546/B 574, A 550/B 578, A 556/B 584); and
- (4) the representation of (1)-(3) as evidence for one's self-conception as an active and spontaneous intellect (from B 158a).

Conjointly these identify any subject who finds them in introspection as a conscious subject, thinking, reasoning and acting in accordance with the same rationality requirements that unify the self. Since those rationality requirements are innate, our capacity to identify these properties of personhood would be similarly innate, or at least very deeply entrenched. Under these circumstances, it would, indeed, be difficult to avoid including these properties in one's self-conception.

Notice the explanatory elegance and simplicity of Kant's account of personhood, under the assumption of the KRT: Structuring our experience according to the subject-predicate relation gives it a basic consistency and coherence that extends to the particular set of contingent empirical concepts thus structured. Satisfying this structural requirement, in turn, is a necessary condition of a rationally unified self. A self that satisfies this rationality requirement thereby generates the cognitively inevitable concept of a reasoning and acting person, which it then applies, first to its own first-personal representations of unified rational agency; and second to those external empirical objects whose behavior exhibit similar adherence to rationality requirements. The concept of rational personhood thereby supplies simultaneously the principles of cognitive organization, self-identification, and recognition of other rational persons in Kant's system. To be a person is to be a self-consciously rational and unified self that manifests its rationality in action.²⁴

Recalling the interpretation of the relation between transcendental and empirical concepts as one of instantiation, the proposal would then be that the transcendental concept of a unified rational person is instantiated by particular empirical exemplars of personhood, among whom each of us necessarily and first-personally counts herself, and inferentially and third-personally counts others. We each necessarily conceive of ourselves as persons, and then use this concept as a criterion for identifying others similarly.

V. Xenophobia

In what follows I will use the terms *person* and *personality* to denote particular, empirical instantiations of personhood as analyzed above. These terms correspond closely to the non-technical use of the terms. Thus when we refer to someone as a person, we ordinarily mean to denote at the very least a social being whom we presume – as Kant did – to have consciousness, thought, rationality, and agency. The term "person" used in this way also finds its way into jurisprudence, where we conceive of a person as a rational individual who can be held legally and morally accountable for her actions. Relative to these related usages, an individual who lacks to a significant degree the capacities to reason, plan for the future, detect causal and logical relations among events, or control action according to principles applied more or less consistently from one occasion to the next is ascribed diminished responsibility for her actions, and her social and legal status as a person is diminished accordingly.

Similarly, when we call someone a "bad person," we communicate a cluster of evaluations that include, for example, assessing her conscious motives as corrupt or untrustworthy, her rationality as deployed for maleficent ends, and her actions as harmful. And when we say that someone has a "good personality" or a "difficult personality," we mean that the person's consciousness, thought, rationality, and agency are manifested in pleasing or displeasing or bewildering ways that are particular to that individual. We do not ordinarily assess a being who lacks any one of these components of personhood in terms of their personality at all. Persons, then, express their transcendent personhood in their empirical personalities.

With these stipulations in place, I now turn to an analysis of the concept of xenophobia based on the foregoing interpretation of Kant. Xenophobia is not simply an indiscriminate fear of strangers in general: it does not include, for example, fear of relatives or neighbors whom one happens not to have met. It is more specific than that. Xenophobia is a fear of individuals who look or behave differently than those one is accustomed to. It is a fear of what is experientially unfamiliar, of individuals who do not conform to one's empirical assumptions about what other people are like, how they behave or how they look. Ultimately it is a fear of individuals who violate one's empirical conception of persons and so one's self-conception. So xenophobia is an alarm reaction to a threat to the rational coherence of the self, a threat in the form of an anomalous other who transgresses one's preconceptions

about people. It is a paradigm example of reacting self-protectively to anomalous data that violates one's internally consistent conceptual scheme.

Recall that on the KRT, if we cannot make sense of such data in terms of those familiar concepts, we cannot register it as an experience at all. I have argued elsewhere²⁵ that *pseudorationality* is an attempt to make sense of such data under duress, i.e. to preserve the internal rational coherence of the self, when we are baldly confronted by anomaly but are not yet prepared to revise or jettison our conceptual scheme accordingly. It is in the attempt to make sense of anomalous data in terms of empirically inadequate concepts that the mechanisms of pseudorationality – rationalization, dissociation and denial – kick in to secure self-preservation. But they succeed in preserving only the appearance of rational coherence. In *rationalization*, we misapply a concept to a particular by distorting its scope, magnifying the properties of the thing that instantiate the concept, and minimizing those that fail to do so. So, for example, conceiving of a slave imported from Africa as three-fifths of a person results from magnifying the properties that appear to support this diminished concept of personhood – the slave's environmental and psychological disorientation, lack of mastery of a foreign language, lack of familiarity with local social customs, incompetence at unfamiliar tasks, depressive physical lethargy (in response to exile and isolation), inarticulateness (in response to the suppression of her culture), childishness (in response to infantilization), resentment and low self-esteem (in response to subjugation and objectification), etc.; and minimizing the properties that disconfirm it – her capacity to learn, to forge innovative modes of communication and expression, to adapt and flourish in an alien social environment, to survive enslavement and transcend violations of her person, etc. In *dissociation*, we identify something in terms of the negation of the concepts that articulate our theory: Identifying Jews as subhuman, blacks as childlike, women as irrational, gays as perverts, or working class people as animals, for example, conceives of them as lacking essential properties of personhood, and so are ways of defining these groups of individuals out of our empirical conceptions of people. In *denial*, we suppress recognition of the anomalous particular or property altogether, by ignoring it or suppressing it from awareness. For example, ignoring a woman's verbal contributions to a conference, or passing over a black person's intellectual achievements when compiling a bibliography, or forgetting to make provisions at a

Christmas celebration for someone who is a practicing Jew are all ways of eradicating the anomalous other from one's domain of awareness.

Thus xenophobia engenders various forms of discriminatory stereotyping – racism, sexism, anti-Semitism, homophobia, class elitism – through the pseudorational mechanisms of rationalization and dissociation, by reducing the complex singularity of the other to a set of oversimplified but manageable properties that invariably diminish our full conception of personhood. For the xenophobe, this results in a provincial self-conception and conception of the world, from which significant available data are excluded – data the inclusion of which would significantly alter the scope and content of the theory. And this provincial theory is sustained with the aid of denial, by enforcing those stereotypes through such tactics as exclusion, ostracism, scapegoating, tribalism, and segregation in housing, education or employment.²⁶ My thesis is that xenophobia is the originating phenomenon to which each of these forms of discriminatory stereotyping is a response. The phenomenon of xenophobia is a special case of a perfectly general human intellectual disposition to literal self-preservation, i.e. preservation of the internal rational coherence and integrity of the self against anomalous data that threaten it.

Nevertheless, to say this much is not to answer the question of how deeply entrenched xenophobia is in our cognitive scheme. Even if it is true that we are innately cognitively disposed to respond to any conceptual and experiential anomaly in this way, it does not follow that our necessarily limited empirical conception of people must be so limited and provincial so as to invite it. A person could be so cosmopolitan and intimately familiar with the full range of human variety that only *The Alien* would rattle her. On the other hand, her empirical conception of people might be so limited that any variation in race, nationality, gender, sexual preference, or class would be cause for panic. How easily one's empirical conception of people is violated is one index of the scope of one's xenophobia; how central and pervasive it is in one's personality is another. In what follows I will want to focus primarily on cases midpoint between such extremes: for example, of a white person who is thoughtful, well-rounded and well-read about the problems of racism in the United States, but who nevertheless feels fearful at being alone in the house with a black television repairman. In all such cases, the range of individuals in fact identifiable as persons is larger than the range of individuals to

whom one's empirical conception of people apply. In all such cases, I will argue, xenophobia can be understood in terms of certain corrigible cognitive errors, only the last of which constitutes full-blown xenophobia.

A. The Error of Confusing People with Personhood

Xenophobia is fueled by a perfectly general condition of subjective consciousness, namely the same first- / third-person asymmetry that, as we saw in Section III, led Kant to propose rational action as a basis for inferring another's personhood. Although I must identify myself as a person because of my necessary, enduring first-personal experience of rationally unified selfhood, my experience of you as a person, necessarily lacking that first-personal experience, can have no such necessity about it:

Identity of person is ... in my own consciousness unfailingly to be found. But when I view myself from the standpoint of another (as object of his outer intuition), this external observer considers me first and foremost in time So from the I, which accompanies all representations at all times in my consciousness, and indeed with full identity, whether he immediately concedes it, he will not yet conclude the objective continuity of my self. For because the time in which the observer situates me is not the same as that time to be found in my own, but rather in his sensibility, similarly the identity that is necessarily bound up with my consciousness, is not therefore bound up with his, i.e. with the outer intuition of my subject. (A 362-363)

Kant is saying that the temporal continuity I invariably find in my own consciousness is not matched by any corresponding temporal continuity I might be supposed to have as the object of someone else's consciousness. Since I am not always present to another as I am to myself, I may appear discontinuously to her consciousness in a way I cannot to my own. And similarly, another may appear discontinuously to my consciousness in a way I cannot to my own.

This is one example of how it can happen, on a Kantian conception of the self, that a necessary concept is instantiated by contingent ones: Although personhood is a necessary concept of mine, whether or not any other empirical individual instantiates it is itself, from my point of view, a contingent matter of fact – as is the concept of that particular individual herself. Though you may exhibit rationality in your behavior, I may not know that, or fail to notice it, or fail to understand it. Nor

can you be a necessary feature of my experience, since I might I ignore or overlook you, or simply fail to have any contact with you. In any of these cases, you will fail to instantiate my concept of personhood in a way I never can. Because the pattern of your behavior is not a necessary and permanent, familiar concomitant of my subjectivity in the way my own unified consciousness and ratiocinative processes are, I may escape your personhood in a way that I cannot escape my own. For me the transcendent idea of personhood is also a transcendental concept that applies necessarily to me, but, from my perspective, only contingently and empirically to you.

Hence just as our empirical experience of the natural world is limited relative to the all-inclusive, transcendent idea of its independent unity, similarly our empirical experience of other persons is limited relative to our all-inclusive, transcendent idea of personhood. But there is an important disanalogy between them that turns on the problem of other minds and the first- / third-person asymmetry earlier described. For any empirical experience of the natural world we have, we must, according to Kant, be able to subsume it under the transcendent concept of a unified system of nature of which it is a part, even if we do not know what that system might be. By contrast, it is not necessarily the case that for any empirical experience of other people we have, we must be able to subsume them under the transcendent idea of personhood. This is because although they may, in fact, manifest their personhood in their personality, we may not be able fully to discern their personhood through its empirical manifestations, if those manifestations fall outside our empirical conception of what people are like.

Suppose, for example, that within my subculture, speech is used to seek confirmation and promote bonding, whereas in yours it is used to protect independence and win status;²⁷ and that our only interpersonal contact occurs when you come to fix my TV. I attempt to engage you in conversation about what is wrong with my TV, to which you react with a lengthy lecture. To you I appear dependent and mechanically incompetent, while to me you appear logorrheic and socially inappropriate. Each of us perceives the other as deficient in some characteristic of rationality: you perceive me as lacking in autonomy and basic mechanical skills, whereas I perceive you as lacking in verbal control and basic social skills. To the extent that this perceived deficit is not corrected by further contact and fuller information, each of us will perceive the other as less of a full-fledged person

because of it. This is the kind of perception that contributes to one-dimensional stereotypes, for example of women as flighty and incompetent or of men as aggressive and barbaric, which poison the expectations and behavior of each toward the other accordingly.

Or take another example, in which the verbal convention in my subculture is to disclose pain and offer solace, whereas in yours it is to suppress pain and advert to impersonal topics; and that our only interpersonal contact occurs when I come to work as your housemaid. Again each of us perceives the other as deficient in some characteristic of rationality: you perceive me as dull and phlegmatic in my lack of responsiveness to the impersonal topics you raise for discussion, whereas I perceive you as almost schizophrenically dissociated from the painful realities that confront us. Again, unless this perceived deficit is corrected by further contact and fuller information, each of us will perceive the other as less of a person because of it, thereby contributing to one-dimensional stereotypes of, for example, blacks as stupid, or of whites as ignorant and out of touch with reality, that similarly poison both the expectations and the behavior of each toward the other.

In such cases there are multiple sources of empirical error. The first one is our respective failures to distinguish between the possession of rationality as an active capacity in general, and particular empirical uses or instantiations of it under a given set of circumstances and for a given set of ends. Because your particular behavior and ends strike me as irrational, I surmise that you must be irrational. Here the error consists in equating the particular set of empirical behaviors and ends with which I am familiar from my own and similar cases with unified rational agency in general. It is as though I assume that the only rational agents there are are the particular people I identify as such. Kant might put the point by saying that each of us has conflated her empirically limited conception of people with the transcendent concept of personhood.

B. The Error of Assuming Privileged Access to the Noumenal Self

But now suppose we each recognize at least the intentionality of the other's behavior, if not its rationality. Since each of us equates rational agency in general exclusively with the motives and actions of her own subculture in particular, each also believes that the motives and ends that guide the other's actions – and therefore the evidence of conformity to the rule and order of rationality – nevertheless remain

inaccessible in a way we each believe our own motives and ends not to be inaccessible to ourselves. This third-personal opacity yields the distinction between the appearance and the reality of the self: You, it seems, are an appearance to me behind which is hidden the reality of your motives and intentions, whereas I am not similarly an appearance that hides my own from myself. The less familiar you are to me, the more hidden your motives and intentions will seem, and the less benevolent I will assume them to be.

Of course whom we happen to recognize as familiar determines whose motives are cause for suspicion and whose are not. There is no necessary connection between actual differences in physical or psychological properties between oneself and another, and the epistemic inscrutability we ascribe to someone we regard as anomalous. It is required only that the other seem anomalous relative to our familiar subculture, however cosmopolitan that may be, in order to generate doubts and questions about what it is that makes her tick. Stereotypes of women as enigmatic or of Asians as inscrutable or of blacks as evasive all express the underlying fear of the impenetrability of the other's motives. And someone who conceives of Jews as crafty, blacks as shiftless, or women as devious expresses particularly clearly the suspicion and fear of various third-personal others as mendacious manipulators that is consequent on falsely regarding them as more epistemically inaccessible to one than one is to oneself.²⁸

Thus our mutual failure to identify the other as a person of the same status as oneself is compounded by scepticism based on the belief that each of us has the privileged access to her own personhood that demonstrates directly and first-personally what personhood really is. The inaccessibility and unfamiliarity of the other's conception of her own motives to our consciousness of her may seem conclusive justification for our reflexive fear and suspicion as to whether her motives can be trusted at all.

Now we have already seen in Section IV that Kant thinks the belief in privileged access is erroneous. From the first-personal relation I bear to my empirical self-conception which I lack to yours, it does not follow that my actual, noumenal motives are any more accessible to me than yours are. Therefore, regardless of how comfortable and familiar my own motives may seem to me, it does not follow that I can know that my own motives are innocuous whereas yours are not. In fact it is difficult to imagine how I might gain any understanding of the

malevolent motives I reflexively ascribe to you at all, without having first experienced them in myself. Of course this is not to say that I cannot understand what it means to be the victim of maleficent *events* without having caused them myself. But it is to say that I must derive my understanding of the malevolent *intentionality* I ascribe to you from my own first-hand experience of it. Therefore your epistemic opacity to me furnishes no evidence for my reflexive ascription to you of malevolent or untrustworthy motives, although that ascription itself does furnish evidence for a similar ascription of them to myself. Thus Kant might put this second error by saying that we have been fooled by the first-/third-person asymmetry into treating the ever-present "dear self" as a source of genuine self-knowledge on the basis of which we make even faultier and more damaging assumptions about the other.

C. The Error of Failing Rationally to Conceive Other Minds

These two errors are interconnected with a third one, namely our respective failures to imagine each other's behavior as animated by the same elements of personhood that animate our own, i.e. consciousness, thought, and rationality. Our prior failure to recognize the other's behavior as manifesting evidence of these properties – a failure compounded by conceptual confusion and misascription of motives – then further undermines our ability to bridge the first-/third-person asymmetry by imagining the other to have them. Since, from each of our first-personal perspectives, familiar empirical evidence for the presence of these properties is lacking in the other, we have no basis on which to make the ascription, and so no basis for imagining what it must be like from the other's perspective. Our respective, limited empirical conceptions of people, then, itself the consequence of ignorance of others who are thereby viewed as different, delimit our capacity for empathy. This is part of what is involved in the phenomenon feminists refer to as objectification, and what sometimes leads men to describe some women as self-absorbed. Kant might put this point by saying that by failing to detect in the other's behavior the rule and order of rationality that guides it, we fail to surmise or imagine the other's motives and intentions.

This error, of failing to conceive the other as similarly animated by the psychological dispositions of personhood, is not without deleterious consequences for the xenophobe herself. Elsewhere I have described the self-centered and

narrowly concrete view of the world that results from the failure to imagine empathically another's inner states, and its interpersonal consequences.²⁹ From the first-personal perspective, this error compounds the seeming depopulation of the social environment of persons and its repopulation by impenetrable and irrational aliens. This is to conceive one's social world as inhabited by enigmatic and unpredictable disruptions to its stability, to conjure chimaeras of perpetual unease and anxiety into social existence. Relative to such a conception, segregation (assuming no relations of interdependence preclude it) is no more effective in banishing the threat than is leaving on the nightlight to banish ghosts, since both threats arise from the same source. Vigilance and a readiness to defend oneself against the hostile unknown may become such intimately familiar and constitutive habits of personality that even they may come to seem necessary requisites of personhood.

D. The Error of Equating Personality with Personhood

The three foregoing errors involve cognitive failures for which a well-intentioned individual could correct. For example, someone who regularly confuses people with personhood might simply take a moment to formulate a general principle of rational behavior that both applies to all the instances with which she is familiar from her particular community and has broader application as well; and remind herself, when confronted by anomalous behavior, to at least try to detect the operation of that principle within it. Similarly, it does not require excessive humility on the part of a person who falsely assumes privileged access to the noumenal self to remind herself that our beliefs about our own motives, feelings, and actions are exceedingly fallible and regularly disconfirmed; and that it is therefore even more presumptuous to suppose any authority about someone else's. Nor is it psychologically impossible to gather information about others' inner states - through research, appreciation of the arts, or direct questioning and careful listening, so as to cultivate one's imaginative and empathic capacities to envision other minds.

Thus it is possible for someone to exhibit these failures without being a xenophobe, just in case she has no personal investment in the defective empirical conception of people that results. A person has a *personal investment* in a conception or theory if

- (1) that theory is a source of personal satisfaction or security to her;

- (2) to revise or reject it would elicit in her feelings of dejection, deprivation or anxiety; and
- (3) these feelings are to be explained by her identification with this theory.

She *identifies with* this theory to the extent that she is disposed to identify it as personally meaningful or valuable to her.³⁰ A person could make the first three cognitive errors without taking any satisfaction in her provincial conception of people ("Is this really all there is?" she might think to herself about the inhabitants of her small town), without identifying with it (she might find them boring and feel ashamed to have to count herself among them), and without feeling the slightest reluctance to enlarge and revise it through travel or exploration or research.

What distinguishes a xenophobe is her personal investment in her provincial conception of people. Her sense of self-preservation requires her conception to be viridical, and is threatened when it is disconfirmed. She exults in the thought that only the people she knows and is familiar with (whites, blacks, WASPs, Jews, residents of Waco, Texas, members of the club, etc.) are persons in the full, honorific sense. This is the thought that motivates the imposition of pseudorational stereotypes, both on those who confirm it and those who do not.

To impose a *stereotype* on someone is to view her as embodying a limited set of properties falsely taken to be exclusive, definitive, and paradigmatic of a certain kind of individual. I will say that a stereotype

- (a) equates one contingent and limited set of valued properties that may characterize persons under certain circumstances with the universal concept of personhood;
- (b) restricts that set to exclude divergent properties of personhood from it;
- (c) withholds from those who violate its restrictions the essential properties of personhood; and
- (d) ascribes to them the essential, disvalued properties of deviance from it.

Thus a stereotype identifies as persons those and only those who manifest the valued properties in the set ((a) and (b)) – call this set the *honorific stereotype* – and subsidiary ones consistent with it (such as minor personality quirks or mildly idiosyncratic personal tastes). And reciprocally, the honorific stereotype by

implication identifies as deviant all those who manifest any properties regarded as inconsistent with it ((c) and (d)) – call this second set of disvalued properties the *derogatory stereotype*. So, for example, an individual who bears all the valued properties of the honorific stereotype as required by (a) may be nevertheless disqualified for membership according to (b), by bearing additional disvalued ones as well – being related by blood or marriage to a Jew, for example; or having bisexual inclinations; or, in the case of a black person, an enthusiasm for classical scholarship. In virtue of violating (b), one may then fail to qualify as a full-fledged person at all (c), and therefore may be designated as deviant by the derogatory stereotype according to (d). The derogatory stereotype most broadly includes all the disvalued properties that fall outside the set defining the honorific stereotype (i.e. "us versus them"), or may sort those properties into more specific subsets according to the range of individuals available for sorting.

A stereotype generally is therefore distinguishable from an inductive generalization by its provincialism, its oversimplification, and its rigid imperviousness to the complicating details of singularity. Perhaps most importantly, a stereotype is distinguishable from an inductive generalization by its function. The function of an inductive generalization is to guide further research, and this requires epistemic alertness and sensitivity to the possibility of confirming or disconfirming evidence in order to make use of it. An inductive generalization is no less a generalization for that: it would not, for example, require working class blacks living in the Deep South during the 60s to dismantle the functionally accurate and protective generalization that white people are dangerous. What would make this an inductive generalization rather than a stereotype is that it would not preclude recognition of a white person who is safe if one should appear. By contrast, the function of a stereotype is to render further research unnecessary. If the generalization that white people are dangerous were a stereotype, adopting it would make it cognitively impossible to detect any white people who were not.

Thus Kant might describe the reciprocal imposition of stereotypes as the fallacy of equating a partial and conditional series of empirical appearances of persons with the absolute and unconditioned idea of personhood that conceptually unifies them. Whereas the first error – of confusing one's empirical conception of people with the transcendent concept of personhood – involves thinking that the only persons there are the people one knows, this fourth error – of equating

personality with personhood – involves thinking that the kind of persons one knows are all there can ever be. So unlike inductive generalizations, the taxonomic categories of a stereotype are closed sets that fundamentally require the binary operation of sorting individuals into those who fall within them and those who do not.³¹

As a consequence of her personal investment in an honorific stereotypical conception of persons, a xenophobe has a personal investment in an honorific stereotypical self-conception. This means that that self-conception is a source of personal satisfaction or security to her; that to revise or disconfirm it would elicit in her feelings of dejection, deprivation or anxiety; and that these feelings are to be explained by her identification with this self-conception. In order to maintain her honorific self-conception, a xenophobe must perform the taxonomic binary sorting operation not only on particular groups of ethnic or gendered others, but on everyone, including herself. Since her self-conception as a person requires her and other *bona fide* persons to dress, talk, look, act, and think in certain highly specific and regimented ways in order to qualify for the honorific stereotype, everyone is subject to scrutiny in terms of it.

This is not only prejudicial to someone who violates these requirements and thereby earns the label of the derogatory stereotype. It is also prejudicial to someone who satisfies them, just in case there is more to her personality than the honorific stereotype encompasses and more than it permits. Avoidance of the negative social consequences of violating the honorific stereotype – ostracism, condemnation, punishment, or obliteration – necessitates stunting or flattening one's personality in order to conform to it (for example, by eschewing football or nightclubs and learning instead to enjoy scholarly lectures as a form of entertainment because one is given to understand that that is the sort of thing real academics typically do for fun); or bifurcating one's personality into that part which can survive social scrutiny and that "deviant" part which cannot (as, for example, certain government officials have done who deplore and condemn homosexuality publicly on the one hand while engaging in it privately on the other). One reason it is important not to equate personality with personhood is so that the former properties can flourish without fear that the latter title will be revoked.

Truncating one's personality in order to conform to an honorific stereotype in turn damages the xenophobe's self-esteem and also her capacity for self-knowledge.

Someone who is deeply personally invested in the honorific stereotype but fails fully to conform to it (as everyone must, of course) views herself as inherently defective. She is naturally beset by feelings of failure, inferiority, shame and worthlessness which poison her relations with others in familiar ways: competitiveness, dishonesty, defensiveness, envy, furtiveness, insecurity, hostility, and self-aggrandizement are just a few of the vices that figure prominently in her interpersonal interactions. But if these feelings and traits are equally antithetical to her honorific stereotype, then they, too, threaten her honorific stereotypical self-conception and so are susceptible to pseudorational denial, dissociation or rationalization. For example, a xenophobe might be blindly unaware of how blatantly she advertises these feelings and traits in her behavior; or she might dissociate them as mere peccadilloes, unimportant eccentricities that detract nothing from the top-drawer person she essentially is. Or she might acknowledge them but rationalize them as natural expressions of a Nietzschean, *übermenschliche* ethic justified by her superior place in life. Such pseudorational habits of thought reinforce even more strongly her personal investment in the honorific stereotype that necessitated them, and in the xenophobic conception of others that complements it. This fuels a vicious downward spiral of self-hatred and hatred of anomalous others from which it is difficult for the xenophobe to escape. Thus the personal disadvantage of xenophobia is not just that the xenophobe devolves into an uninteresting and malevolent person. She damages herself for the sake of her honorific stereotype, and stunts her capacity for insight and personal growth as well.

A sign that a person's self-conception is formed by an honorific stereotype is that revelation of the deviant, disvalued properties provokes shame and denial, rather than a reformulation of that self-conception in such a way as to accommodate them. For example, a family that honorifically conceives itself as white Anglo-Saxon Protestant may deny that its most recent offspring in fact has woolly hair or a broad nose. Similarly, a sign that a person's conception of another is formed by a derogatory stereotype is that revelation of the other's non-deviant, valued properties provokes hostility and denial, rather than the corresponding revision of that conception of the other in such a way as to accommodate them. For example, a community of men that honorifically conceives itself in terms of its intellectual ability may dismiss each manifestation of a woman's comparable intellectual ability as a fluke.³²

These two reactions are reciprocal expressions of the same dispositions in the first- and third-personal cases respectively. Shame involves the pain of feeling publicly exposed as defective, and denial is the psychological antidote to such exposure: for example, if the purportedly WASP offspring does not have negroid features, there is nothing for the family to feel ashamed of. So a person whose self-conception is defined by an honorific stereotype will feel shame at having disvalued properties that deviate from it, and will attempt to deny their existence to herself and to others. By contrast, hostility toward another's excellence is caused by shame at one's own defectiveness, and denial of the excellence is the social antidote to such shame: for example, if the woman is not as intelligent as the men are purported to be, then there is no cause for feeling shamed by her, and so none for hostility toward her. So a person whose self-conception is formed by an honorific stereotype will feel hostility toward a derogatorily stereotyped other who manifests valued properties that violate that derogatory stereotype; and will attempt to deny the existence of those valued properties in the other to herself and to others.

In the first-personal case, the objects of shame are disvalued properties that deviate from one's honorific stereotypical self-conception. In the third-personal case, the objects of hostility are valued properties that deviate from one's derogatory stereotypical conception of the other. But in both cases the point of the reactions is the same: to defend one's stereotypical self-conception against attack, both by first-personal deviations from it and by third-personal deviations from the reciprocal stereotypes this requires imposing on others. And in both cases, the reactions are motivated in the same way: the properties regarded as anomalous relative to the stereotype in question are experienced by the xenophobe as an assault on the rational coherence of her theory of the world – and so, according to Kant, on the rational coherence of her self.

Indeed, left untreated, all four of these cognitive errors more generally – the conflation of the transcendent concept of personhood with one's provincial conception of people that another happens to violate, the ascription to the other of malevolent motives on the basis of an epistemically unreliable self-conception, the inability to imagine the other as animated by familiar or recognizably rational motives, and the equation of personality with personhood inherent in the imposition of reciprocal stereotypes – combine to form a conception of the other as an inscrutable and malevolent anomaly that threatens that theory of the world which

unifies one's experience and structures one's expectations about oneself and other people. If this were an accurate representation of others who are different, it would be no wonder that xenophobes feared them.

VI. Xenophilia

Now recall once more Kant's original claim about the structure of the self (the KRT). He said that if a perception failed to conform to the categories of thought that unified and structured the self, it could not be experienced by that self at all. Also recall that we detected an ambiguity in Kant's claim: it was unclear whether a perception would have to conform to (A) both the transcendental and the empirical concepts that unified the self, or (B) only the transcendental ones, in order to be minimally an object of experience. Suppose (A) is correct, and perceptions must conform both to the transcendental and to the empirical concepts that structure the self and its experience. Then these sets of concepts are materially equivalent: something is a person if and only if it falls under one's empirical conception of people. Then someone must conform not only to my transcendental concept of personhood, but also to my empirically contingent and limited concept of what persons are like – i.e. of people – in order for me to recognize her as a person. Therefore, (A) implies that an anomalous other who violates my limited conception of people thereby violates my transcendental conception of personhood as well.

We have already seen in Section III that even if the concept of personhood is transcendental as well as transcendent according to Kant, this concept is at best an instantiation of the transcendental substance-property relational category. Since my transcendental concept of personhood is not equivalent to the transcendental concept of a thing or substance in general, my failure to recognize the other's personhood does not imply a failure to recognize her as an object with properties altogether. I may recognize another who is anomalous with respect to my concept of personhood as consistent with my concept of objects in general. However, if the other must conform to my limited conception of people in order to conform to my concept of personhood but does not, then from my perspective, an object is all that she can ever be. In this case, xenophobia is a hard-wired cognitive disposition that is impervious to empirical modification.

But suppose instead that (B) furnishes the correct account of the relation between transcendental and empirical concepts, such that perceptions need conform

only to the transcendental concepts and not necessarily to the empirical ones, in order to be part of one's coherent experience. (B) leaves open the possibility that a person might have an empirically limited conception of people yet fail to be a xenophobe, just in case she acknowledges as a matter of principle that there must be other ways to do things and other ways to live besides those with which she is familiar; and just in case she is able to put this principle into practice when confronted by some of them. This is the case described in Section V. D, of the individual who commits cognitive errors A-C, but has no personal investment in the defective empirical conception that results.

(B) also leaves open the possibility that one could be a xenophobe in the sense discussed in Section V.D, yet be corrigible in one's xenophobia. For (B) acknowledges the possibility that even though the xenophobe equates her limited conception of people with her transcendental concept of personhood, someone might conform to her transcendental concept of personhood without conforming to her empirical conception of people. That is, in this case it is cognitively possible to introduce into her range of conscious experience a new object the behavior of which satisfies the rule and order of rationality even though it fails to satisfy her honorific stereotype of personhood. And it is possible for her to recognize in this conceptually anomalous behavior the rule and order of rationality, and so the personhood of another who nevertheless violates that honorific stereotype.

Since recognition of the existence of such an anomaly constitutes a counterexample to her honorific stereotype of personhood, the xenophobe has two options according to (B). Either she may, through the mechanisms of pseudorationality, seek some strategy for explaining this anomaly away; or else she may revise her stereotypic and limited conception of people in order to accommodate it. Thus (B) suggests that it is in theory possible for the xenophobe to reformulate and reform that conception in light of new data that disconfirms it, and so to bring her reciprocal stereotypes closer to open-ended inductive generalizations.

Of course whether or not this occurs, and the extent to which it occurs, depends on the virulence of her xenophobia; and this, in turn, on the extent of her personal investment in her honorific, stereotypical self-conception. But if (B) is correct, and one can discern the personhood of someone who violates one's limited conception of people, then pseudorational dismissal of the stranger as a person is not a viable option. By hypothesis the properties that constitute her identity as a person

cannot be denied. Attempts to dissociate them, i.e. to dismiss them as insignificant, alien or without value have unacceptable implications for one's own which similarly must be pseudorationalized out of the picture.³³ Moreover, attempts to rationalize them as flukes or mutations or illusions or exceptions to a rule undermine the universality of the rule itself. As in all such cases, pseudorationality does not, in fact, preserve the rational coherence of the self, but only the appearance of coherence in one's self-conception, by temporarily dismissing the anomaly that threatens it. In the event that a xenophobe is confronted with such a phenomenon, xenophobia conflicts with the requirements of literal self-preservation and finally must be sacrificed to it. So finally, the only way for the xenophobe to insure literal self-preservation against the intrusion of an anomalous person is to revise her reciprocal stereotypes of herself and others accordingly so as to integrate her.

There is evidence in the text of the first *Critique* that supports (B) as Kant's preferred alternative. These are in those introductory, explicative sections of the *Dialectic*, in which Kant maintains that it is in the very nature of transcendent concepts of reason to have a breadth of scope that surpasses any set or series of empirical experiences we may have; indeed, to provide the simplest unifying principle for all of them and more. Thus, for example, he tells us that "the principle peculiar to reason in general, in its logical use, is: to find for the conditioned cognitions of the understanding the unconditioned whereby its unity is brought to completion." (A 307/B 364) By the "conditioned," Kant means those experiences and rules that depend on an inferential relation to other, more inclusive principles that explain them. And by the "unconditioned," Kant means those principles, concepts or ideas of reason that are not themselves dependent on any further ones but rather provide the explanation of all of them. What he is saying here is that rationality works interrogatively for us: given some datum of experience we understand, we reflexively seek to enlarge our understanding by searching for further data by which to explain it.

Kant then goes on to say in the same passage that this logical principle becomes a transcendent one through our assumption that if dependent explanatory rules and experiences are given, then the whole series of them, ordered in relations of subsumption of the sort that characterize a covering-law theory, must be given as well; and that this series is not itself dependent on any further explanatory principles.³⁴ Kant's point is that we assume that any limited explanation of

experience we have is merely part of a series of such explanations that increase in generality and inclusiveness, up to a maximally inclusive explanation of all of them. Thus we regard each such partial experience of the world we have as one among many, all of which are unified by some higher-level theory. And later he says that

[t]he transcendental concept of reason is none other than that of proceeding from a totality of conditions to a given conditioned. Now since only the unconditioned makes the totality of conditions possible, and conversely the totality of the conditions is itself always unconditioned; so a pure concept of reason in general can be explained through the concept of the unconditioned, so far as it contains a basis of the synthesis of the condition. (A 322/B 379) ... concepts of pure reason ... view all experiential knowledge as determined through an absolute totality of conditions. (A 327/B 384; also see A 311/B 368, B 383-385, A 409, A 509)

What he means is that we regard any particular phenomenon as embedded in a systematically unified series of such phenomena, such that if we can explain some partial series of that kind, then there is an entire series of which that partial series is a part that we can also explain; and such that that more inclusive explanation explains everything there is about the phenomenon to explain. So Kant is saying that built into the canons of rationality that structure our experience is an inherent disposition to seek out all the phenomena that demand an inclusive explanation, and to test its inclusiveness against the range of phenomena we find.

These remarks support (B) because they imply that the innate cognitive concepts that structure and unify our experience invariably, necessarily outstrip our empirical conceptions of it. Kant is saying that it is in the nature of our cognitive limitations – i.e. that we can only have knowledge of sense-based experience – that the explanatory scope of the innate concepts that structure and unify it necessarily exceeds that sensory basis itself. This means that we view any experience in implicit relation to other possible experiences of its kind, and finally in relation to some systematic explanation that makes sense of all of them. So no single experience, or series of experiences, can ultimately satisfy our appetite for conceptual completeness, because the scope of the higher-level concepts we invoke to explain them necessarily outstrips the limited number of those experiences themselves. There will always be a lack of fit between our innate rational capacity and the empirical theories it generates, because they will always appear limited in scope in a

way our innate capacity for explanation itself does not. So no matter how much sensory data we accumulate in support of our empirical theories of ourselves or the world, we are so constructed intellectually as to be disposed to feel somewhat dissatisfied, inquisitive, restless about whether there might not be more to explain, and to search further for whatever our search turns up.³⁵

But this means that we are disposed reflexively to regard anomalous data as more than mere threats to the integrity of our conceptions of the world and ourselves, for the disposition to inquire further and to seek a more inclusive explanation of experience remains, even when literal self-preservation has been achieved. We also are disposed to regard those data as irresistible cognitive challenges to the scope of our conceptions, and as provocations to reformulate them so as to increase their explanatory reach. Because, according to Kant, we are always seeking the final data needed to complete the series of experiences our conceptions are formulated conclusively to explain, it could even be said that we are disposed actively to welcome anomalies, as tests of the adequacy of the conceptions we have already formulated.

When applied specifically to the transcendent idea of personhood, this disposition to welcome anomaly as a means of extending our understanding amounts to a kind of xenophilia. That is, it amounts to a positive valuation of human difference as intrinsically interesting and therefore worthy of regard, and a disvaluation of conformity to one's honorific stereotypes as intrinsically uninteresting. It dismantles the assumption that there is any cause for self-congratulation or self-esteem in conforming to any stereotype at all, and represents anomalous others as opportunities for psychological growth rather than mere threats to psychological integrity. It implies an attitude of inquiry and curiosity rather than fear or suspicion, of receptivity rather than resistance toward others; and a belief that there is everything to be gained, and nothing to be protected, from exploration of another person's singularity.³⁶ We often see this belief expressed in the behavior of very young children, who touch, poke, prod, probe and question one without inhibition, as though in knowledge of another there were nothing to fear. What they are lacking, it seems, is contingent empirical evidence to the contrary.

Footnotes

¹ Work on this paper was supported by an NEH Summer Stipend and a Woodrow Wilson International Scholars' Fellowship. Portions are excerpted from Chapter XII of a manuscript in progress, *Rationality and the Structure of the Self*. It has benefited from presentation to the Wellesley Philosophy Department Faculty Seminar, and also from the comments of Anita Allen, Alison MacIntyre, John Pittman, and Kenneth Winkler.

²Immanuel Kant, *Kritik der Reinen Vernunft*, herausg. Raymund Schmidt (Hamburg: Felix Meiner Verlag, 1976). All references to this work are parenthecized in the text. Translations from the German are my own. Connoisseurs will find my translations to be generally more literal than Kemp-Smith's, and (I think) more accurate in conveying not only the substance of Kant's claims, but his manner of expression. Despite Kant's tendency to indulge in run-on sentences, he is by and large a plain speaker with a fondness for the vernacular, not the stilted, pretentious Prussian Kemp-Smith makes him out to be. But the major objection to Kemp-Smith's translation is that he obscures important philosophical issues by overinterpreting Kant so as to resolve them before the monolingual English reader can become aware that there is anything to dispute. This is particularly evident in the debate about transcendental content (see below, Section II and Footnotes 16 and 17).

³P. F. Strawson, *The Bounds of Sense* (London: Methuen, 1968).

⁴Immanuel Kant, *Grundlegung zur Metaphysik der Sitten*, herausg. Karl Vorlander (Hamburg: Felix Meiner Verlag, 1965). All references to the Academy Edition are parenthecized in the text. Translations from the German are my own.

⁵John Rawls, *A Theory of Justice* (Cambridge, Mass.: Harvard University Press, 1971), Chapters III and VII; Alan Gewirth, *Reason and Morality* (Chicago: University of Chicago Press, 1978), Chapter II; Thomas Nagel, *The Possibility of Altruism* (Oxford: Oxford University Press, 1975), Part Two. I discuss Rawls' recent transition away from a Humean model of rationality in "Personal Continuity and Instrumental Rationality in Rawls' Theory of Justice," *Social Theory and Practice* 13, 1 ((Spring 1987), 49-76, and Nagel's and Gewirth's reliance on a Humean model of motivation in Chapters II and III respectively of my *Rationality and the Structure of the Self*.

⁶ My exegetical remarks in this paper should not be mistaken for a defense of the extended overall interpretation of Kant they clearly presuppose. I defend this interpretation against the canonical views in *Kant's Metaethics*, in progress.

⁷This thesis may be viewed as the resolution of a Gedankenexperiment Kant earlier conducts at A 89-91, in which he entertains the possibility of unsynthesized appearance. In any case, his ultimate commitment to this thesis is clear. See Robert Paul Wolff, *Kant's Theory of Mental Activity* (Cambridge, Mass.: Harvard University Press, 1968) for a discussion.

⁸See, for example, P. F. Strawson, *ibid.*, Chapter II.2. In hindsight Kant himself grudgingly admits that hypothetical and disjunctive syllogisms contain the same "matter" as the categorical judgment, but refuses to budge on their essential difference in form and function. See Kant's *Logic*, trans. Robert Hartman and Wolfgang Schwarz (New York: Bobbs-Merrill, 1974), Paragraphs 24-29, 60, Note 2., especially pages 111 and 127.

⁹This thesis is elaborated in the contemporary context by Gerald M. Edelman, *Neural Darwinism: The Theory of Neuronal Group Selection* (New York: Basic Books, 1987) and *The Remembered Present: A Biological Theory of Consciousness* (New York: Basic Books, 1989). See the review of Edelman and others by Oliver Sacks in "Neurology and the Soul," *The New York Review of Books* XXXVII, 18 (November 22, 1990), 44-50.

¹⁰"Rationality and the Structure of the Self," excerpted from *Rationality and the Structure of the Self* and delivered to the Association for the Philosophy of the Unconscious, American Philosophical Association Eastern Division Convention, Boston, Mass., 1986.

¹¹*ibid.*

¹²Thomas Kuhn, *The Structure of Scientific Revolutions* (Chicago: University of Chicago, 1971), Chapters VI-VIII.

¹³"Two Conceptions of the Self," *Philosophical Studies* 48, 2 (September 1985), 173-197, reprinted in *The Philosopher's Annual VIII* (1985), 222-246; also see "Pseudorationality," in Amelie O. Rorty and Brian McLoughlin, Eds. *Perspectives on Self-Deception* (Los Angeles: University of California Press, 1988), 297-323.

¹⁴"Higher-Order Discrimination," in Amelie O. Rorty and Owen Flanagan, Eds. *Identity, Character and Morality* (Cambridge, Mass.: MIT Press, 1990), 285-309; reprinted in condensed form in the monograph series, *Studies on Ethics in Society* (Kalamazoo, Mich.: Western Michigan University, 1990).

¹⁵Kant definitively identifies human desires and inclinations as empirical concepts at A 15/B 29. In the following sections I will offer an interpretation of the relation between transcendental and empirical concepts as one of instantiation. This would treat the empirical concepts of desires and inclinations as instantiations of the transcendental concept of passion, and the empirical concept of intentional human behavior as an instantiation of the transcendental concept of action.

¹⁶The German sentence runs as follows: <<Da die Zeit nur die Form der Anschauung, mithin der Gegenstände, als Erscheinungen, ist, so ist das, was an diesen der Empfindung entspricht, die transzendente Materie aller Gegenstände, als Dinge an sich (die Sachheit, Realität).>>

¹⁷Kant's statement here of course makes a great deal of trouble for his doctrine of transcendental idealism and therefore is not developed significantly in the first *Critique*. The Refutation of Idealism, for instance, provides no conclusive evidence either of his acceptance or rejection of such a view. However, there are other passages and problems in the first *Critique* and *Prolegomena* that furnish evidence of Kant's underlying commitment to it.

Specifically, the view that the sufficient condition for the correct application of empirical concepts is given by the transcendental matter of things in themselves answers a question regarding the status of what Kant entitles the "matter of appearance" that remains unanswered through both editions of the first *Critique*. In both editions, Kant clearly wants to say that the form of appearance is spatiotemporal intuition, which inheres innately in the transcendental subject and is empirically real. And in both editions he contrasts the form of appearance with its matter, which is "that in the appearance which corresponds to sensation" and is given *a posteriori* (A 20/B 34). But exactly where and to what the sensation is given, and what exactly is the nature of the correspondence, remains unclear. Kant defines sensation as "the effect of an object upon the faculty of representation, so far as we are affected by it." (A 19/B 34) Here he is clearly referring to the transcendental subject's faculty of representation. And since empirical objects are the consequence

of that subject's cognitive activity, they cannot be supposed to exert causal influence on it. So by "an object" (*Gegenstand*), Kant must mean a non-empirical object, i.e. a thing in itself. So he is claiming that there is a non-empirical object which, by affecting the transcendental subject's faculty of representation, causes that subject to feel sensations.

Kant also denies that there is any "subjective representation, referring to something *outer*, which could be called objective *a priori*." (A 28/B 44) So, in particular, sensation (a subjective representation which, by corresponding to the matter of appearance, presumably refers to it) cannot be empirically real. Therefore, although sensation is the causal effect of an object on the transcendental subject, and although it refers to something outer, namely the matter of appearance, it is also an *a posteriori* subjective representation which, unlike intuition, is not empirically real.

If sensation is *a posteriori*, one would expect to find it in the empirical world of appearance, and this is exactly where Kant locates it in the A Edition. There Kant's justification for assigning this status to sensation is that tastes, for example, belong to the "special constitution of sense in the subject that tastes it," and colors similarly are not properties of the objects we see, "but only modifications of the sense of sight, which is affected in a certain manner by light." By contrast with space, which is a necessary part of appearances, "[t]aste and colors are ... connected with the appearances only as effects accidentally added by the particular constitution of the sense organs. ... grounded in sensation, and, indeed, in the case of taste, even upon feeling (pleasure and pain) as an effect of sensation." (A 28-29) So colors are effects of sensation, and taste is an effect of feeling, which in turn is an effect of sensation. And what is it that affects the sense organs, so as to give rise to the sensation that in turn causes one to perceive, say, colors? Kant tells us that it is light.

But light is itself an appearance, just as the sense organs are among the appearances of the empirical self and not part of the transcendental subject to whom the empirical self appears. So the secondary qualities of appearances such as color and taste must result from the effect of some of those appearances, such as light, on other appearances, such as the empirical self's sense organs. This explains how the empirical self comes to experience the secondary qualities of appearances: it experiences them as sensory effects of empirical appearances on its sense organs. But the mere spatiotemporal *form* of an appearance cannot be supposed to have such

causal efficacy. If anything about an appearance does, it must be its *matter*. So the matter of appearances cannot be supposed to be identical to the secondary qualities it may cause the subject to generate. So although these appearances have matter, they do not have secondary qualities except in so far as these are ascribed to them by a sensing empirical subject. Thus when Kant describes sensation as "corresponding" to the matter of appearance, he seems to be suggesting a three-place causal relation: the matter of appearance causes the empirical subject's sensation, which in turn causes the empirical subject to perceive the secondary qualities she ascribes to it.

Locating sensations in causal relations between the empirical self and the natural world of objects accords well with empirical psychology. The problem is that without explaining the connection between empirical sense organs and transcendental sensibility, this account obscures Kant's claim that sensations are the effect of a nonempirical object on the representational ability of the transcendental self. For since empirical objects are the product of transcendental cognitive activity, they cannot themselves engender the activity that produces them. And since things in themselves are supposed to be beyond our cognitive capacity to understand, Kant is not entitled to assert their effect on the subject, either.

What Kant should do is break his own rule of silence on what things in themselves can and cannot do, just this once. He should say that sense organs may be the way transcendental sensibility appears to the introspecting transcendental subject (see Section IV, below), just as empirical objects such as light may be the way things in themselves appear to that subject. That way the "*transcendental* matter of things in themselves" (A 143/B 182) could causally affect the subject's sensibility such that it then generated sensations, and so the secondary qualities of empirical objects. This would make sense of the sentence at A 143/B 182 that Kemp-Smith mistranslates. It would stipulate a simpler, two-place causal relation between sensation and the transcendental matter that corresponds to it, namely that the latter causes the former.

Moreover, it would make sense of Kant's claim at A 20/B 34 about the matter of appearance corresponding to sensation. For if appearances could be the way things in themselves appear to the transcendental subject, and sensations occur in the transcendental subject, which itself appears as the empirical subject, then one way for both the matter of appearance and the transcendental matter of things in

themselves to correspond to sensation would be if these two kinds of matter were, so to speak, materially equivalent. In this case, the causal relation of transcendental matter to the transcendental subject's sensations would appear as a causal relation between the matter of appearance and the empirical subject's sense organs.

Another benefit of this interpretation for the present discussion would be that it would supply detailed support for my suggestion, immediately following, that an instantiation relation between transcendental and empirical concepts preserves relevant content from systematically related things in themselves, through the transcendental and finally empirical concepts that structure that content. The material equivalence of transcendental and empirical matter would offer some evidence of what that content might be, and how increasingly specific conceptualizations of cognitively available particulars might preserve it.

This interpretation would require Kant only to revise the doctrine of transcendental idealism, not necessarily to abandon it. In particular, it would require him to revise his claim in the A Paralogisms, that

in fact, when one regards outer appearances as representations, which are effected in us by their objects as existing things in themselves outside us, it is not possible to see how one can know their existence otherwise than through the inference from the effect to the cause, relative to which it must always remain doubtful whether the latter is in us or outside us. One can, indeed, concede that there may be something which is, in the transcendental sense, outside us and is the cause of our outer intuitions, *but this is not the object that we understand under the representations of matter and corporeal things; for these are merely appearances, i.e. mere types of representations, which are to be found only in us, and whose reality is based on immediate consciousness, just as is the consciousness of my own thoughts.* (A 372; italics added)

Consistent application of Kant's strictures about the unknowability of things in themselves would require Kant to replace the italicized passage, which makes a positive, substantive claim about what things in themselves cannot be, with one that admits our inability to *know whether or not* the "something which is in the transcendental sense outside us and is the cause of our outer intuitions" is "the object that we understand under the representations of matter and corporeal things." Kant is entitled to say that such an object is at least an appearance to be found in us. He is

also entitled to say that, by hypothesis, we can know nothing about such object beyond its appearance to us. But he is not entitled to deny that it might, in fact, accurately represent the nature of things in themselves as well.

This interpretation would, however, require Kant to jettison his allegiance to the traditional distinction between primary and secondary qualities. Since both would now be generated *a priori* by the subject's innate faculty of sensibility, and both would refer to something "outer" – respectively, the form and matter of appearance, the asymmetry between them would be far less striking. But Kant could still maintain that secondary qualities vary from person to person (i.e. to the extent that the transcendental subject's senses do) whereas primary ones do not; and so continue to insist that secondary qualities were not, unlike space and time, empirically real.

In the B Edition Kant moves closer to such a view. Here his argument for denying that sensations are empirically real is that "they belong merely to the subjective constitution of our *manner of sensibility*, for instance, of sight, hearing, touch, as in the case of the sensations of colors, sounds, and heat, which, since they are mere sensations ... do not of themselves yield knowledge of any object ..." (A 28/B 44; italics added) On the next page Kant goes on to deny that colors and tastes can "be rightly regarded as properties of things, but only as changes in the subject, changes which may, indeed, be different for different people. ... [and] ... with reference to color, can appear differently to every eye." (B 45; the passage in German runs, << ... jedem Auge in Ansehung der Farbe anders erscheinen kann.>> Kemp-Smith *again* willfully mistranslates this passage as "... in respect of its color, can appear differently to every observer."). In this version of the argument Kant ascribes the five senses, and eyes in particular, to the subject's transcendental sensibility. Clearly this is a strategic error: Kant should not ascribe apparent properties of empirical objects, such as the sense organs human beings happen to have, to the transcendental subject to whom these properties empirically appear. But the interest of this *faux pas* is the evidence it provides of Kant's actual view. It clearly implies that he does think transcendental subjects have senses even if he shouldn't say so. And it supports the above suggestion that these senses can be understood as appearing empirically as sense organs to empirical observation or introspection.

What remains is to provide at least some textual evidence that the transcendental subject's sensations are caused by things in themselves. We have just seen that sensations cannot come from the empirical objects that are their consequences, as the A Edition suggests. Either they are self-generated by the subject or they come from something else. Now Kant insists that sensibility is a purely receptive capacity for "receiving representations through the way in which we are affected by objects." (A 19/B 33) Moreover, he later identifies and discusses that class of representations he thinks are actively generated by the mind's effect on itself (B 68-69; see Section IV, below). So we can infer that he doesn't think sensations can be actively self-generated. If they come from something else, this can only be from things in themselves. The above remarks, imploring Kant to speak up about the behavior of things in themselves, offers a possible account of this behavior that would square nicely, not only with A 143/B 182, but with A 19/B 34 as well.

Now for Kant actually to state this in the *Critique* would constitute a commitment to causal realism that conflicted with his strictures that we can know nothing of things in themselves, and in particular cannot assert the applicability of the categories to them. So Kant refrains from any such claim in the first *Critique*. Luckily for us, by the time he writes the *Prolegomena* he is ready to tip his hand. There Kant states quite clearly, in contrasting his own view with that of the idealist:

I, on the contrary, say that things as objects of our senses existing outside us are given, but we know nothing of what they may be in themselves, knowing only their appearances, *that is, the representations which they cause in us by affecting our senses. Consequently I grant by all means that there are bodies without us, that is, things which, though quite unknown to us as to what they are in themselves, we yet know by the representations which their influence on our sensibility procures us.* (Immanuel Kant, *Prolegomena to Any Future Metaphysic*, trans. Lewis White Beck (New York: Bobbs-Merrill, 1950), Ac. 288, italics added)

Surely Kant's considered commitment to causal realism, achieved with intellectual and temporal distance from the many ambiguities and confusions of the first *Critique*, to things in themselves as causal sources of the appearances they effect in us by impinging on our senses, could not be any clearer than this.

¹⁸Scientific paradigm shifts needn't invalidate Kant's insight, since a gavagai doesn't stop being a gavagai when we discover that it is "really" a perturbation in the electromagnetic field. I discuss the requirement of inclusiveness at greater length in "Seeing Things," *The Southern Journal of Philosophy XXI*, Supplementary Volume (1990): Moral Epistemology, 29-60.

¹⁹"Pseudorationality," *op. cit.* Note 13.

²⁰The use of *bestimmen* can also mean to decree or ordain someone to do something; but it cannot mean merely to cause something. *Bestimmen* always carries the connotation of shaping some idea or event by cognitive means. Therefore it does rule out "cause" as a synonym for "determine." Kant's usual words for causality are *Kausalität* or *Ursache*. So *Selbstbestimmung* would refer to the cognitive activity of resolving to be or act in a certain way, not to that of merely causing oneself to do so. The tendency to think of Kant's concept of self-determination on analogy with that of causal determination should be resisted at all costs.

²¹ I develop this claim at greater length in "The Meaning of 'Ought' and the Loss of Innocence," Invited Address delivered to the American Philosophical Association Eastern Division Convention, Atlanta, Ga., December 1989. Abstracted in the *Proceedings of the American Philosophical Association* 63, 2 (October 1989), 53-54.

²²*op. cit.* Note 13.

²³Lest I be charged with the same fault I charge Kemp-Smith, I note the ambiguity of the German: <<Sie, die Vernunft, is allen Handlungen des Menschen in allen Zeitumständen gegenwärtig und einerlei, selbst aber ist sie nicht in der Zeit, und gerät etwa in einen neuen Zustand, darin sie vorher nicht war; etc. >> [italics added] A literal translation of this passage would make the meaning incoherent, so I infer that Kant was expressing himself ungrammatically.

²⁴Now in the *Groundwork*, Kant claims that "the human being and in general every rational being *exists* as end in himself, *not merely as means* for arbitrary use by this or that will, but rather must be viewed as *at the same time an end* in all of his actions, whether directed to himself or to other rational beings. ... rational beings are called *persons* because their nature distinguishes them as ends in themselves, i.e. as something that must not be used merely as means, and thus so far restricts all power of choice (and is an object of respect)." (Ac. 428) Besides an immediately preceding paragraph that introduces these concepts with definitions of them, there is little in

the *Groundwork* of an explicit nature to have prepared the reader for these remarks on personhood as an end in itself, so it may seem that Kant has simply pulled these intuitively appealing ideas out of a hat. Moreover, Kant does not explain in the *Groundwork* why it is that personhood or rational nature deserves to be regarded as an end in itself, or even what he thinks an end in itself is.

The explanations of Kant's claims lie, rather, in the first *Critique*. There Kant characterizes an end as a species of idea (A 318/B 375). As we have seen, an *idea* is for Kant a technical term that denotes a final outcome of the rational disposition to generalize inclusively from lower-level to higher-level concepts, principles, and theories. He also describes the peculiar sphere of reason as an order of ends which is at the same time an order of nature, and human beings as the only creatures in nature who can contain the final end of this order in themselves and also exempt themselves from it through morality (B 425). So the sphere of rationality is one in which all of our experience is systematically organized and unified according to inclusive theoretical concepts in the manner already discussed. Human beings both contain the final end of this natural order within themselves and also can transcend it through moral conduct. In the Canon of Pure Reason, Kant tells us what this final end is: it is the idea of a natural world made moral, in which the free power of choice of rational beings "has, under moral laws, thoroughgoing systematic unity as such, as much with itself as with the freedom of every other." (A 808/B 836) In this moral world, the supreme good is happiness as directly apportioned to moral worth by a Supreme Reason that rules according to moral law, and we are rationally compelled to envision this world as the outcome of our efforts to achieve moral worthiness to be happy (A 809/B 837 - A 811/B 839, A 813/ B 841 - A 816/B 844). The ultimate end, Kant tells us, is the entire vocation of man, and this is treated by moral philosophy. (A 840/B 868) So Kant says in the *Groundwork* that personhood is an end in itself because a person has the capacity rationally to represent to herself, as a final end of her moral conduct, a divinely just moral order in which she participates as an equal member. This is the same vision that lies behind Kant's obscure remarks about membership and lordship in the kingdom of ends (Ac. 433-434).

²⁵*Op. cit.* Note 13.

²⁶*Op. cit.* Note 14.

²⁷This is the main thesis of Professor Deborah Tannen's fascinating *You Just Don't Understand: Women and Men in Conversation* (New York: William Morrow and Co., Inc., 1990), a popularization of her research in linguistics on gender differences in language use.

²⁸I chart the systematic use of such disvaluative properties in "Higher-Order Discrimination," *op. cit.* Note 14.

²⁹"Impartiality, Compassion and Modal Imagination," *Ethics* 101, 4: Symposium on Impartiality (July 1991), 726-757.

³⁰The concept of personal investment is discussed in my "Moral Theory and Moral Alienation," *The Journal of Philosophy* LXXXIV, 2 (February 1987), 102-118. Also see *op. cit.* Notes 13 and 14.

³¹I am indebted to Rüdiger Bittner for pressing this question in discussion of my "Higher-Order Discrimination," although I was unable to address it properly in that context.

³²*op. cit.* Note 14 for a fuller discussion.

³³A case study of this phenomenon might be the Post-Modernist attitude of mourning over the loss of value and meaning in contemporary creative and intellectual products of "Western civilization" at just that historical moment when the longstanding contributions to it by women and people of color are gaining recognition.

³⁴I argue that Kant's moral theory is a descriptive, explanatory theory that fits the deductive-nomological model in "The Meaning of 'Ought' and the Loss of Innocence," *op. cit.* Note 21.

³⁵This idea of theoretical rationality and theory-building as an innate disposition is given some support by Robin Horton's cross-cultural work. See his "African Traditional Thought and Western Science," in *Rationality*, Ed. Bryan Wilson (Evanston, Ill.: Harper and Row, 1970), 131-171. As I understand Horton's conclusions, the main difference between Western scientific theories and the cosmologies of traditional societies is that the latter lack the concept of modality, i.e. recognition of the conceptual possibility that the favored and deeply entrenched explanation may not be the right one or the best one. They therefore lack the attitude of epistemic uncertainty that leads in the West to the joint problems of

scepticism and solipsism. To this extent the stance of intellectual dissatisfaction I am attributing to Kant's epistemology may be culturally specific.

³⁶Thus xenophilia in the sense I am defining it should be distinguished from a superficially similar, but in fact deeply perverse form of xenophobia, in which the xenophobe reinforces her honorific, stereotypical self-conception by treating the other as an exotic object of research, whom (like a rare species of insect) it is permissible to examine and dissect from a superior vantage-point of inviolate disingenuity. By contrast, the xenophile acknowledges the disruption and threat to the integrity of the self caused by the other's difference, and seeks understanding of the other as a way of understanding and transcending the limitations of her own self-conception.