Towards A Research Program in Kantian Positive Psychology
DRAFT

In recent years, Kantians have begun to take up the challenge of responding to philosophical appropriations of empirical psychology that seem to undermine key claims in Kant's moral theory. In particular, Pauline Kleingeld (2014, 2015) and Patrick Frierson (2010, 2013, under-review) have, in different ways, responded to situationist critiques of the possibility of character. While not from a specifically Kantian perspective, Selim Berker (2009) has responded to Joshua Greene's now famous article entitled "The Secret Joke of Kant's Soul," (2008), in which Greene uses contemporary neuroscience to argue that we should prefer our consequentialist moral reasoning to our deontological moral intuitions. In addition, there has been a steady increase of attention to what Robert Louden famously called "Kant's Impure Ethics" (Louden 2000). In the context of this focus on empirically-informed features of Kant's philosophy, there has been some attention to working out the ways that Kant saw his empirical science of human behavior – "pragmatic anthropology" – as different both from other eighteenth century approaches to empirical psychology and from other sciences within Kant's system (see especially Sturm 2009, Cohen 2009). So far, however, there has been little recent attention to thinking through what it would mean to carry out a Kantian research program in psychology today. In this paper, I use a brief summary of how Kant might respond to situationist research in psychology as a springboard for discussing problems and proposals for developing a distinctively Kantian research program in empirical psychology.

1. Situationism and the Challenge to Kant's Moral Philosophy

One of the most prominent recent appropriations of psychological research into moral philosophy in recent years has been the "situationist" challenge to virtue ethics. Situationists such as John Doris, Gilbert Harman, and Mark Alfano draw on psychological studies showing that human beings behave quite differently in different situations, even when not consciously aware of causally-efficacious situational variables: "The experimental record suggests that situational factors are often better predictors of behavior than personal factors . . . To put it crudely, people typically lack character" (Doris 2002:2). The supposed relevance to moral theory is that because "virtue ethics presupposes that there are character traits of the relevant sort, that people differ in what character traits they have, and these traits help to explain differences in the way people behave," virtue ethics is empirically false (Harman 2000:168). Thus, "[r]ather than striving to develop characters that will determine our behavior in ways substantially independent of circumstance, we should invest more of our energies attending to the features of our environment that influence behavioral outcomes" (Doris 2002:146).

While situationist criticisms of "virtue ethics" are generally directed towards broadly Aristotelian approaches to ethical theory, Kant's ethical theory is at least prima facie susceptible to situationist critique because the Kantian good will requires a firm *commitment* to moral principles. It is not enough to act properly or even to act properly for good reasons. One must act properly from good reasons to which one is *consistently* – that is, as a matter of "character" – committed (see 6:63; 25:648; 25:1176). *This* sort of good will seems to depend upon a conception of character threatened by the results of situationist research.

However, these threats are merely apparent, and situationist empirical research, properly understood and interpreted, actually supports Kant, properly understood and interpreted. In two other papers, I have laid out the details of this Kantian response to situationism (xxx-deleted for

review). The core element of that response is what I call "moral pessimism" (xxx-deleted for review). Whereas philosophers like Doris and Alfano believe that people are generally good or at least not very bad, Kant's view is that human beings are radically evil. "Character is rare" (7:292) in part because evil is universal (6:32). Evidence that human beings generally lack character only confirms Kant's argument for radical evil; it does not undermine his moral theory.

Moreover, the fact that humans *do* lack character cannot, for Kant, imply that humans *must* lack character, and that for three reasons. First, empirical sciences in general study only the "empirical character" of things, the way that they appear to human beings given our particular forms of cognition. But the character for which we are ultimately morally responsible is an "intelligible character" that serves as the ground of whatever empirical character one turns out to have. Second, unlike physics, psychology has no a priori basis that could provide strictly necessary psychological laws akin to Newton's laws. Psychology offers at best a natural history of the soul, such that any "lack of character" discovered by psychologists is an empirical generalization, not a necessary truth. Finally, Kant's moral philosophy establishes *on a priori grounds* that character is morally required. We ought to act on maxims that can be universalized, where "maxims" already involve sustained commitment to practical principles, and the relevant "universalization" precludes volitional inconsistency (see especially 6:24-5). Because ought implies can (5:30), consistency must be possible.[1]

This response to situationists like Doris and Harman deals with empirical facts about the situation-dependence of human action by treating them as evidence of radical evil in human beings rather than as reasons for altering moral theory. Such a response might make it seem as though, for Kant, empirical research into the determinants of human behavior is worthless from the standpoint of ethics, or useful only for establishing that humans are evil, or at best valuable for cataloging the different sorts of propensity to evil in human beings. This approach could justify a sphere of potentially vibrant albeit narrowly-focused Kantian studies oriented towards showing the limited scope of situationist research. One model for this kind of work would be Lawrence Kohlberg and his students' work on Milgram's experiments, showing shared features of those agents who resisted the situational pressures in those experiments.[2] Another model would be recent studies showing how one's "mindset" can reduce or even eliminate the impact of various situational variables, such as the framing effects of racial labels.[3] By articulating clear even if only approximate criteria for distinguishing between those who seemingly have Kantian character and those who do not, a Kantian research program could specifically investigate whether and to what extent such character is resistant to various situational pressures. Insofar as it is – and Kohlberg, Dweck, and others give reasons to think that it might be – this would show (problematic) situation-dependence to be a contingent and preventable trait, rather than a fixed aspect of human nature.

In fact, however, Kant's response to situationism goes further. It opens the possibility for a renewed and better-focused use of empirical psychology within ethics. If, as Kant insists, the *good will* – that is, a will that has good *character* – is the only thing good without qualification, then the most important features of humans' environments are those that help cultivate wills less susceptible to accidental environmental influence. In order for empirical psychology to live up to its full ethical potential, psychologists need not only to diagnose the problem – that humans lack character – but to investigate methods for *solving* this problem. That is, a *Kantian* empirical psychology would be

> moral anthropology, which would deal … with the subjective conditions in human nature
> that hinder people or help them in fulfilling the laws of a metaphysics of morals. It would

deal with the development, spreading, and strengthening of moral principles (in education in schools and in popular instruction) and with other similar teachings and precepts based in experience. (6:217)

This sort of Kantian psychology would investigate what can be done – either directly or indirectly – to cultivate Kantian character. It would provide a distinctively Kantian version of relatively young and vibrant new research programs in "positive psychology" and related character studies.

In this paper, I briefly lay out the core components of Kantian positive psychology, along the lines of the moral anthropology described in the quotation above. My focus is on the broad outlines of such an anthropology rather than specific principles, experimental procedures, or the like. While I stick relatively close to Kant's own accounts of moral anthropology, my ultimate goal is not to defend his particular empirical claims but to outline the overall framework of his approach.

2. Moral Philosophy and the Value of Psychological Research

As an empirically-rooted discipline investigating the conditions for strengthening moral principles, Kantian moral anthropology essentially depends upon a *prior* understanding of what the relevant moral principles are. Kant opens his *Groundwork* by carefully distinguishing empirical investigation from a priori investigation and describing his own moral philosophy as, first and foremost, "a pure moral philosophy, completely cleansed of everything that may be only empirical and that belongs to anthropology" (4:389). In that context, he insists that we can take from experience neither criteria for morality nor even determinate limits to human moral possibilities (4:406). Kant's warnings about "practical men (for whom morals is mere theory)" who "pretend to see in advance, from the nature of the human being, that *he is* never *going to will* what is required" and thereby "*make* improvement *impossible*" (PP 8:371, 373, cf. TP 8:276f.) anticipate related warnings by psychologists such as Carol Dweck, who notes that self-concepts that allow for growth and improvement are invaluable for actually growing and developing (Dweck 2006), or Lisa Shu and her colleagues, who note that

bad behavior motivates moral leniency and leads to the strategic forgetting of moral rules … [P]eople could set off on a downward spiral of having ever more lenient ethics and even more unethical behavior. (Shu et. al. 2011:344, cited in Miller 2014:79)

Relatedly Mark Alfano's positive endorsement of virtue-ascriptions as "self-fulfilling prophecies" (Alfano 2013:9) has a negative correlate in situationism itself. Precisely insofar as situationists convince people that they are not responsible or not *capable* of virtue, they actually increase the likelihood that people will be incapable of it. One of the main ways in which humans' tendency to self-deception works to foster and cultivate radical evil in human nature is precisely through watering down the demands of morality and/or weakening our responsibility for our own actions (and characters). Situationism does both. Recent appropriations of situationism lead to precisely those moral theories of "lenient justice" (27:331) that have the effect of making human beings more like what they describes us as being, and conjoined with an increased emphasis on good results rather than good characters, human beings will tend to be even *more* situation-driven, gradually moving from frail or impure wills that are situation-dependent despite endorsing the importance of genuine virtue towards *depraved* wills that settle for mere legality rather than genuine morality, a life of doing good things, but without a good *will*.[4] This point is confirmed in recent studies (e.g. Kohlberg and Candee 1984; Dweck 2000, 2006) showing that one of the main factors that makes people susceptible to sub-rational influences is the belief that they lack control or responsibility for their behavior or attributes. It

is no surprise that Milgram's final exhortation to test subjects – "You have no choice; you must go on" – involves specifically seeking to undermine their sense that they have control over their own decisions.

This Kantian criticism of situationism depends upon a particular conception of what the purpose of moral theory is. If one assumes that the purpose of ethical theory is to promote behavior with good effects in the world, then the attempt to promote good character could be a danger. As Doris insists, "Many times a confidence in their character is precisely what puts people at risk in morally dangerous situations" (147). Situationist psychology provides some evidence that one can best promote good behaviors and effects through manipulating situations in which people find themselves. And many social psychologists see moral improvement primarily in terms of just such consequences. In a revealing insight into the moral preconceptions of social psychologists, Aharon Kleinberger aptly remarks,

> Since philosophers have an occupational bias in favor of reflective morality, there are hardly any philosophic theories of ethics which hold overt behavior in accordance with accepted moral norms to be a sufficient condition for the morality of an act or an agent. But common sense tends to make moral judgments according to such an implicit conception. Not only the man in the street seems to adopt this position but so do many social and behavioral scientists. (Kleinberger 1982:149, quoted in Kohlberg and Candee 1984:55)

We might add philosophers like Doris and Harman to the chorus of "men in the street" and "scientists" who see morality as primarily about overt behavior and who thus see any attempt to cultivate internal states of character, except where these manifest in such overt behavior, as a waste of moral energy. As Doris explains, "ethical reflection is in the business of helping people *behave* better" (Doris 2002:166, emphasis added). Only in such contexts can the results of situationism be taken to imply, as Doris takes them to imply, that "[r]ather than striving to develop characters that will determine our behavior in ways substantially independent of circumstance, we should invest more of our energies attending to the features of our environment that influence behavioral outcomes" (Doris 2002:146).

If good *behavior* is *not* the highest good, however, promoting such behavior through behavioral manipulation may turn out to do more "harm" – in a morally loaded sense of harm – than good. In terms of Kant's ideal of the good will, manipulation oriented towards overt behavior most obviously fails to cultivate good wills, since the moral status of the will is a matter of its stable commitment to good maxims. But such manipulation is also directly objectionable from a moral point of view in that it treats human agents merely as means to good effects. Provoking "morally" desirable behavior by means of situational manipulation bypasses rather than respects that capacity for self-determination that is the essence of the dignified humanity worthy of Kantian respect (see 4:428-9).

In short, moral theory determines the value of psychological research. *Kant's* moral theory, with its emphasis on action out of a respect for the moral law that is an enduring part of one's character, requires a research program oriented towards *cultivating* such enduring respect rather than one that excuses or manipulates overt behaviors. Psychological research thus should discover techniques for bringing about the true end of morality: consistently good wills, that is, human beings who have a settled practical commitment to both the general principle of willing only in accordance with duty (understood in terms of something like the categorical imperative) and to specific moral principles that follow from this general one. Central to this moral end is the cultivation of character *as such*, so empirical psychology should devote considerable attention to those cases in which human beings *do* exhibit signs of character, seeking to discern

what developmental conditions and personal traits foster such character.

3. Kant's Positive Psychology

For Kant, then, situationist studies of the rarity of character are important in a negative sense, but far more important is careful attention to ways of *cultivating* character in human beings. And Kant's own moral anthropology offers considerable details about how to accomplish this task. While he claims that some innate natural endowments are either helpful or necessary for the development of character, most of Kant's attention is on contingent aids to the cultivation of character, of which education is the most important: "The acquisition of good character with people happens through education" (25:1172). Not just any education will do, however, so Kant's discussions of character are often accompanied by specific pedagogical recommendations, such as avoiding imitation (5:154; 7:325; 9:487-8; 25:599, 635, 722ff., 1386) and using uplifting moral stories and examples (5:156; 8:286).[5] Kant also mentions further factors such as participation in polite society (see Frierson 2005b), stable and just political regimes, peace, and progress in arts and sciences (8:375). Beyond external influences, Kant offers practical guidance for how one can cultivate character in oneself by following some basic maxims of self-care: "a) Not to speak an untruth intentionally … b) not to dissemble … c) not to break one's legitimate promise … d) not to join the company of evil-minded people … e) not to pay attention to slander" (7:294, cf. 25:1387-88, 1392).

Kant may have been wrong about these specific claims. His justification for the detrimental effects of imitation in education or the salutary effects of polite society are not, unlike his moral theory itself, grounded in a priori claims about what is essential to character (or the good will) as such. They are empirical claims about what forces can strengthen one's ability to act on principles, moral or otherwise, in a consistent way. They are based on careful observation, plausible psychological reconstruction, and some armchair speculation, rather than rigorous experimental study. But they offer conjectures that can and should be tested in a Kantian psychological research program. Kant himself hoped his own work would provide bases for further research, "headings under which this or that observed human quality of practical relevance can be subsumed," by means of which "the growth of science for the common good is promoted and accelerated" (7:121-2). And with respect to pedagogy in particular, which is among the most important influences on cultivating character, Kant not only insisted that "experiments in education are … necessary … and experience teaches that our experiments often show quite different effects from the ones expected" (9:451), but he also directly financially supported the experiments in education of his contemporary Basedow (see 2:445-452). To test Kant's claim that "imitation … greatly hinders character, [so] in education one must never refer one's children to the neighbor's children ... but rather build their character directly [with] principles of good and bad" (25:635), one might compare curricula of moral education focused on comparing children with one another to curricula that emphasize the articulation of moral ideals through principles.[6] To test Kant's claim that taking an aesthetic interest in the beautiful, particularly in nature, "indicates a disposition of mind that is favorable to moral feeling" (5:229), one might run interventions that focus on aesthetic cultivation (perhaps one-off priming interventions but ideally systematic programs of aesthetic cultivation) and then check for the effects of these sorts of interventions on tests of moral integrity, such as the cheating studies of Mazar et. al. (2008). In both cases, one would focus not merely in one-time improvements in moral behavior; rather, as in Good, Aronson, and Inzlicht's 2003 study showing how a growth mindset can reduce stereotype threat effects, the goal would be to test the extent to which various interventions can cultivate a moral character that is resistant to morally irrelevant situational

influences.

While specific examples of Kantian hypotheses would reward experimental study, the more important role of Kant's moral anthropology is to provide paradigms of the *kinds* of psychological theories that should be part of such a research program. These hypotheses focus not on what is possible for human beings in general as one finds them, but on what methods are most effective for helping human beings come to have the qualities – particularly character – that are morally required. While not involving proposals specific to Kant's empirical psychology, Carol Dweck's claims about the effects of praise on developing a resilient growth mindset (see Dweck 2000, 2006), Angela Duckworth's specific techniques for fostering "deliberate practice" (Duckworth 2016), and Edward Deci's descriptions of "social-contextual factors that nurture intrinsic motivation and promote internalization" (Deci et. al. 1999) all involve investigation of ways to cultivate traits of character that make people ultimately less susceptible to irrelevant situational factors. Regardless of which particular empirical hypothesis we choose to investigate, we can glean from Kant's overall approach to anthropology several basic principles for broadly Kantian research programs in empirical psychology. These principles would be necessary conditions for considering psychological research to be pragmatic in a Kantian sense. The rest of this section describes four basic principles of a broadly Kantian framework for empirical psychology.[7] For each principle, I briefly highlight how it might apply to a hypothetical experiment for testing one specific claim of Kant's pragmatic anthropology, that in society, when people

> adopt the illusion of affection, of respect for others, of modesty, and of unselfishness without deceiving anyone at all, because it is understood by everyone that nothing is meant sincerely by this …[,] when human beings play these roles, eventually the virtues, whose illusion they have merely affected for a considerable length of time, will gradually really be aroused and merge into the disposition. (7:151)

This "politeness hypothesis," as I will call it, posits that mere politeness, even when everyone sees it to be an illusion, can eventually cultivate genuine virtue. One way to test this would be simply to look for correlations between politeness and various moral characteristics. A better approach would involve, say, testing the thesis within a school setting, where we might compare a character education program that emphasizes norms of politeness with an otherwise similar program that lacks this emphasis, measuring the effects of each on both polite behaviors and the cultivation of moral character. Here, then, are four principles of Kantian pragmatic psychology:

1. **Moral theory precedes psychological research.** This would be the most important contribution of Kant to empirical psychology today, and one widely ignored by situationists and other philosophers seeking to integrate empirical psychology into normative philosophy (including even Kantians like Habermas (1993)). In terms of the practice of psychology, this principle requires shedding the pretense of non-controversial normative goals and recognizing a priority for philosophical theory in setting parameters for psychological investigation. With respect to the thesis that politeness eventually gives rise to real virtue, any experiment to test this effect must begin with a clear articulation of what "virtue" is, and in particular what differentiates mere polite behavior from genuine virtue. A purely behavioral analysis of virtue could easily distinguish the two in terms of specific types of behavior, such that politeness relates to merely conventional standards of behavior (using "please" and "thank you" or eating with a knife and fork) while virtue is a matter of more universally necessary behaviors (refraining from false promising, offering assistance,

avoiding free-riding). For Kant, by contrast, the key distinguishing feature of virtue is the incorporation of moral principles into one's disposition, as opposed to politeness which involves *mere* behavior. Of course, this conception of virtue – or any moral claim that would lie at the heart of pragmatic empirical psychology – depends fundamentally upon philosophical defense, and Kant's idealism, moral theory, and philosophy of psychology provide great resources for such defense. Moreover, the principle must be interpreted, as Kant did, to refer only to the basic structure and principles of a normative theory; particular applications and even certain basic moral requirements can be modified in the light of empirical work.[8]

2. **Moral standards must not be relaxed due to the psychological difficulties of attaining those standards.** In taking empirical research as a basis for specifying, clarifying, refining, and even partially revising moral (and other normative) requirements, we must follow Kant's example and maintain high standards even in the face of empirical evidence that those standards are difficult to attain. If moral pessimism is justified, then it is justified; we shouldn't pretend that we are better than we really are. More importantly, in psychological research, we should aim for the cultivation of moral excellent *according to the highest standards of such excellence*. Even if this aim requires subsidiary and more attainable goals, those goals should not supplant the moral ideal. In the case of the politeness hypothesis, the claim that virtue is cultivated through polite behavior is *not* a claim that whatever can be so cultivated counts as virtue. It may turn out that polite behavior really cultivates only habits of interaction and not actual virtue (that is, behavior motivated by adherence to moral principles). In this case, it might be tempting to redefine virtue in terms merely of a broader and more consistent scope of respectful behavior towards others. For Kant, however, the politeness thesis is specifically a thesis about the transformation of mere behavioral tendencies into a moral disposition, where the nature of that disposition is defined through moral theory prior to and independent of the results of experimental research.

3. **The only thing good without qualification is the good *person*, not good states of affairs, actions, or states of mind.** Situationist critics of character typically appeal, implicitly or explicitly, to good *actions* or *states of affairs* as the ultimate moral goal. Even defenders of character often describe the goal in terms of whatever psychological traits lead to socially or personally beneficial behaviors. For Kant, however, the only thing good without qualification is a good *will*, so psychology should focus on the means for cultivating such wills. *Personality* and *traits* rather than *behavior* should be the central foci of psychological research, and – as a regulative ideal of pragmatic psychological research – traits must be seen as *cultivatable* rather than fixed (cf. Dweck 2000, 2006). In section five, I will argue that this principle gives rise to problems with developing effective experimental measures of outcomes. Here I note only that in the case of politeness, the independent variable – politeness – can relatively easily be set in terms of patterns of behavior and interaction. However, the dependent variable – virtue itself – is not an action or even pattern of action but rather *virtue*, an inner disposition characterized by strength in "overcoming all sensible impulses that oppose" moral principles (6:397). What the experimental psychologist seeks to measure is the effect of certain patterns of polite behavior on shaping the *person* into someone whose moral fortitude is increasingly dominant over non-moral influences.[9]

4. **Self-conscious and reflective self-control is more important than good habits, feelings, or unconscious motivational states.** Kant's conception of character, and the good will that depends on character, is premised on the idea that what matters morally is what we exert self-

conscious control over. The ultimate justification for this is first-personal and morally-reflective: what we *hold* ourselves responsible for is what we self-consciously control. Research that shows, say, how to refine our unconscious motivations or how to cultivate (mostly unconscious) habits of sensitivity and good behavior is at best *indirectly* moral-anthropological: "virtue is not to be defined … merely as an *aptitude* and … a long-standing *habit*" but must "result from considered, firm, and continually purified principles" for "if the practice of virtue were to become a habit the subject would suffer loss of that *freedom* in adopting his maxims that distinguishes an action done from duty" (6:383, 409). This criterion implies that central elements of moral judgment such as sensitivity to morally salient particulars of situations (see Herman 1993) should not be the *primary* foci of moral education as such. The cultivation of *Kantian* character is not the cultivation of good habits, moral sensitivity to particulars, or other non-reflective dispositional states, and any psychology that focuses on the cultivation of these features misses the most morally important target. Thus experiments that show that polite behaviors gradually become internalized and give rise to unconscious patterns of respectful interaction would *not* confirm the politeness hypothesis. They would show, instead, that politeness threatens to undermine the *sort* of respect that constitutes genuine moral worth.

These four principles would be core elements of a broadly Kantian psychological research program. Methodologically, they specify the respective roles of philosophical argument and empirical research. Substantively, they orient psychologists in terms of the kinds of questions and problems most worthy of research. But they underspecify the scope of any possible research program in several respects. They do not dictate any particular psychological methodology, for example, and they do not involve any particular claims about the physical or physiological bases of psychological characteristics. Kant has opinions about these issues, insisting that observation is superior to experiment for empirical psychological research (see 25:1435) and that such research should not concern itself with possible physiological bases of psychological capacities (see 7:119).[10] Kant offers important and underappreciated arguments for these particular methodological claims, and one might want to describe "Kantian psychology" more narrowly to build in more of Kant's specific claims about the nature of psychology. On my view, however, the four core principles outlined above identify a sufficiently "Kantian" approach to empirical psychology, even without additional elements of his theory. These principles even leave unstated precisely what moral criteria will be used for determining what counts as a good will, such as the role of *principles* of action taken to hold *universally* or the notion of acting "for the sake of duty." A fully determinate research program would need to specify these criteria, and Kant offers excellent arguments for his particular way of doing so, but the four principles above give a basic overview of a broadly Kantian approach to empirical positive psychology.

4. Kant and Contemporary Positive Psychology

In his emphasis on the improvement of normal human lives, Kant's anthropology resembles the burgeoning "positive psychology" movement today, which aims to "catalyze a change in the focus of psychology from preoccupation only with repairing the worst things in life to also building positive qualities" (Seligman and Csikszentmihalyi 2000:5).[11] Positive psychology and the character studies associated with it are still sufficiently young subfields within psychology that the development of real Kantian alternatives to the consequentialist-Aristotelian models currently being employed in positive psychology could have real impact. There is a need for providing better focus even for the existing "results-oriented" and "fact-

based" studies of how to cultivate character. Given widespread public appeal to "psychological facts" to justify social policy and even personal efforts at raising children and self-improvement, there is a demand for psychological research that can improve human lives. But in the absence of good philosophical frameworks, psychologists and thereby the general public risk basing lives and social programs on bad moral philosophy. Kantians, in cooperation with sympathetic psychologists, need to offer good psychology, backed by good philosophy.

For research programs like positive psychology, Kant's contribution could be invaluable. Once one turns from describing human behavior to seeking methods for improving ordinary human lives, one needs a clear standard for what counts as *improvement*. This is particularly important when one is interested not merely in happiness but also in moral development. And while Kant's pragmatic anthropology shares with positive psychology a focus on human well-being in the broadest sense (Kant 10:146, 25:734-35; Seligman 2000, 2005), his *moral* anthropology focuses particularly on moral improvement (6:217). Like positive psychologists, Kant sees "character" improvement as a central component of both well-being and moral development.[12] And positive psychologists themselves highlight the importance – and difficulty – of outlining precisely what sort of "character" is worth developing. Christopher Peterson, one of the leading positive psychologists working on character development, rightly asks,

> [W]hat is character? As long as we fail to identify the specifics, different groups in our society – despite their common concern for human goodness – will simply talk past one another when attempting to address the issue [of how to cultivate character]. (Peterson and Seligman 2004; see too Peterson 2006)

The more they seek to better human lives, the more evident to (positive) psychologists becomes the Kantian point that any valuable pragmatic psychology must orient itself in terms of a conception of the good that *precedes* empirical investigation into the causes that help and hinder the promotion of that good.

Unfortunately, contemporary positive psychology has developed a philosophically unsophisticated approach to defining the nature of character, one driven by the attempt to find articulable and measurable traits that are widely esteemed in a variety of cultures. In a defining investigation into character, leading positive psychologists chose desirable traits based on "historical and cross-cultural reviews" (Peterson and Seligman 2004:8), something like a literature review of great works in philosophy and religion, seeking a non-controversial conception of the good life. Strikingly, their philosophical exploration of conceptions of virtue was guided by Katherine Dahlsgaard – a developmental psychologist – rather than by a trained philosopher or intellectual historian, who would have more readily recognized and emphasized important conflicts amongst various conceptions of the good. The result is a treatment that pretends non-controversial universality and ends up with an extremely thin and nonetheless controversial conception of virtue. By contrast, the first principle of Kantian positive psychology outlined in section three makes explicit that determinate and even controversial moral commitments must precede psychological research, and that one must argue for the legitimacy of one's values. Even insofar as Kant and contemporary psychologists both appeal to moral common sense (or, in Kant's terms, "common rational moral cognition"), positive psychologists have too readily pursued an inductive approach that looks for shared features among folk prejudices rather than a transcendental approach that seeks the essential conditions underlying moral assessment as such.

In part because of the thinness of their philosophical conception of character, many positive psychologists make their concept of virtue (or character, or happiness) more determinate

using standards of empirical tractability rather than philosophical plausibility. That is, starting from relatively vague conceptions of a "family of positive dispositions, characteristics like perspective, teamwork, kindness, and hope" (Peterson 2006:139, cf. Peterson and Seligman 2004, Lopez and Snyder 2009:28f.), they make these concepts more determinate by defining traits in such a way that "explicit criteria" and "assessment devices" are readily available (Peterson and Seligman 2004:7, cf. Seligman 2002:130-31). The determinate content of character is thus provided by the standards of ease of measurement, ignoring both the philosophically controversial standards that their resultant characteristics embody and the possibility that the most relevant characteristics may be extremely difficult to assess empirically. In the case of the politeness hypothesis, this approach risks missing the central phenomenon Kant aims to explore. If virtues are defined by ease of empirical measurement, there is serious danger of taking the politeness that is a mere appearance of virtue to be equivalent to the virtuous disposition that Kant claims can be cultivated by affecting that appearance. The effort to find a conception of character that is easily measurable and as non-controversial as possible fits perfectly within general norms of much contemporary social science. By contrast, as we will see in the next section, Kant's value-laden starting points and emphasis on character raise real problems for empirical psychology, but not ones that should be ignored.

Finally, even within the requirements to be measurable and general, positive psychologists have some leeway about how to understand particular virtues, but generally speaking, they endorse a eudaimonism that sees virtue primarily as a component in overall life-satisfaction, rather than as a distinct and supreme good. In its shift to character, positive psychology also seems to have been influenced by the fact that the dominant philosophical voices advocating the importance of "virtue" and "character" during the period in which positive psychology began were predominantly Aristotelian (see Peterson and Seligman 2004:10, Seligman 2002:112, 130), as well as a general interest in connecting scientific psychology with the popular self-help movement. Thus even within their survey of philosophical and religious approaches, the great books of their cross-cultural comparisons led to an overemphasis on *Classical* big-name philosophical sources (Plato and Aristotle) rather than more recent philosophers. The main texts from which they draw their conception of character and virtue are sources traditionally seen as religious (Taoism, Buddhism, Hinduism, Judaism, Christianity, and Islam) and among the accounts most clearly identified with philosophy, they turn to only Confucius, Plato, and Aristotle (see Dahlsgaard, Peterson, and Seligman 2005). Given the prominence of Plato and Aristotle in defining the philosophical conception of virtue in their accounts, positive psychologists understandably focus on character as a matter of dispositions that need not be self-conscious or particularly reflective and that are more habitual than principle-based.

By contrast, for Kant – as noted in the fourth principle in the previous section – self-conscious reflection is more important than habits of behavior. Thus, in the case of the politeness hypothesis, if a study showed that socially-enforced polite behavior gradually gives rise to unconscious habits of respectful interaction, this would *dis*confirm *Kant's* politeness hypothesis but might be taken as a great *success* for a more Aristotelian positive psychologist. Moreover, character involves controlling oneself in accordance with *principles* rather than ad hoc case-by-case considerations. Kant's specific conceptions of the moral requirements for the good will are at odds with the way that positive psychologists describe the general consensus about the nature of character and virtue, and Kantian philosophers can and should contest those psychologists' conception of character, proposing a Kantian alternative conception of what

"positive" psychology would be.

      This "Kantian positive psychology" would emphasize the cultivation of character as a general tendency to remain self-consciously consistent in one's commitments to self-prescribed principles of action. It would include research in developmental psychology focused on the conditions for the cultivation of character in children, including not only general environmental conditions but also specific sorts of interventions that could facilitate better character development. Examining whether promoting patterns of polite interaction facilitates the development of character would be just one example of such studies. It would also look at social conditions that facilitate or inhibit character formation in adults. And it would emphasize – as contemporary positive psychology does well – techniques that individuals can use for *self*-strengthening of character. The goal in all of these cases is not primarily to discern what conditions make it easier to *act* consistently, but to look for the means by which one can cultivate *wills* that more easily resist the influences of both blatant temptations and what should be irrelevant features of one's situation. Some such studies could be done in the relatively isolated experimental format of much empirical psychology. For examples that are not specifically Kantian but are consistent with the spirit of Kant's pragmatic psychology, we might look to Casey and Burton's 1982 study showing that more generalized moral principles ("be honest") were more effective than more specific ones ("do this" and "don't do that") at getting children to consistently act on the principles they were taught, or to Good et. al.'s work on self-concept and stereotype threat, or Carol Dweck's work on how to prime and cultivate self-concepts oriented towards moral and intellectual growth. The most important studies here would be longitudinal ones focusing on long term impacts of various strategies of pedagogy, social structure, and self-improvement. Because of the cost and difficulty of such studies, it is important both to have a clear sense of the goal at which one aims and realistic expectations for results (e.g., that one should some statistically significant differences with different strategies for character development, rather than the rapid discovery of easy techniques for establishing character in anyone).

5. "Problems" for Kantian empirical psychology

      This article has outlined a Kantian research program in positive psychology and sketched how such a program would relate to other contemporary approaches. Psychological research conducted in accordance with Kantian principles would be a fundamentally different from contemporary situationist research. Rather than looking for situational causes of various behaviors, it would investigate developmental factors that could influence the development of (good) character defined in a particular, Kantian, way. Moreover, as I showed in section one, while situationist research is *different* than the research required by Kant, nothing in situationist research *undermines* the prospects for such a Kantian research program. Nonetheless, there are some distinctive problems for specifically Kantian positive psychology. This section highlights and responds to two important features that differentiate psychology from other sciences in Kant's account, and that make such psychology difficult to bring into line with prevailing contemporary norms of social science research.

      First, while the object of psychology is the empirically-given human being, the ultimate goods that positive psychology seeks to promote are moral goods. Unlike physics and chemistry, any Kantian psychology grounded on or oriented towards normative ideals that are established prior to and independent of empirical psychological research will be philosophically controversial. Granted that physics is grounded in a priori laws of cognition, Kant's a priori

moral principles of freedom play a fundamentally different role in psychology than the a priori conditions of experience play for physics. In Kantian positive psychology, these moral principles set goods to be promoted, rather than constitutive principles guiding theory construction.

In the contemporary context, many scientists – both physicists and psychologists – often eschew commitment to specific moral values, particularly given moral philosophers' disagreements about those values. There is a widespread desire within contemporary empirical psychology to orient research towards either the "fact-based" articulation of appropriate goals or, more commonly, the detailed investigation of means towards achieving goals that researchers pretend to be universally accepted ideals. Thus basic standards of psychological health, vague concepts of "happiness" or "character," or broadly consequentialist social goods are made the ultimate ends, and psychology investigates the means. One important role for Kantian empirical psychology and the philosophy that supports it must be clear articulation of specific ideal ends, but given the prevalence of utilitarian and virtue ethical conceptions of human goods, Kantian ideals will be contested. It may turn out that the means to these Kantian ends are also means to other widely accepted ends; Kant, for instance, saw the establishment of character as both necessary for having a good will and beneficial for having a happy life. But it is not only possible but also likely that many means for cultivating character in Kant's sense will put at risk other possible end-goals, such as maximally beneficial behavior or quick and reliable unconscious responses to situations. In the case of the politeness hypothesis, the specific mechanisms that Kant proposes for explaining how politeness promotes virtue involve an increased self-consciousness about one's ability to control one's behavior in particular circumstances (see Frierson 2005b). If Kant's account of moral virtue is correct, this mechanism – rather than, say, habituation – would be morally beneficial, but if moral virtue ought to be conceived in terms of rapid, situationally-sensitive habits of behavior, this mechanism could actually be a hindrance. If – as seems to me likely – there turn out to be different social configurations that promote different sorts of polite behavior, what best promotes *Kantian* virtue may turn out not to best promote "virtue" understood in other ways. However the empirical research may turn out, essential to a Kantian psychological research program is a willingness to pursue research oriented towards promoting moral (and other) ideals that are sufficiently specific to be philosophically controversial.

Fortunately, there is no good reason that psychological research cannot be oriented in this way. There may be political reasons that governments or foundations choose not to fund research that is self-conscious and explicit about its controversial normative premises, but existing research, while it pretends towards universalism, is just as normatively loaded as Kantian research would be. Positive psychology by definition must work towards the promotion of various value-loaded goods. And given real controversy about such goods, no positive psychological research program can aim towards determinate goals without aiming towards controversial ones. Positive psychology seeks universally agreed upon ideals of character, but the more specific those ideals are, the less universal they are; the appearance of universality is an illusion.[13] In that context, a Kantian positive psychology would have the *advantage* of making its value commitments explicit and in offering philosophical defense of those commitments.

Second, and much more problematically, the emphasis on *persons* rather than states of affairs or deeds makes it hard to empirically verify the success of various means for cultivating character (or the good will). According to Kant, this problem is severe and intractable:

It is absolutely impossible by means of experience to make out with complete certainty a

single case in which the maxim of an action otherwise in conformity with duty rested simply on moral grounds and on the representation of one's duty. (4:407)

With respect to character as such, it can be difficult to distinguish the mere "rigid, inflexible disposition" of a "Charles XII" from genuine *character* (7:293). Even focusing on persons as such makes empirical study more difficult. Behavior and states of affairs can be relatively easily observed, described, recorded, classified, and – ideally – reduced to quantitative measures. It is much more difficult to access the psychological states that give rise to behavior, and even harder – for Kant, impossible in principle – to gain empirical knowledge of the free "persons" that underlie those psychological states. Moreover, psychological states, particularly states of character, are not the sorts of things that can be easily quantified, a point Kant emphasizes as a basis for claiming that "empirical psychology can never rise to the level of a science" (4:471). The ultimate inscrutability of humans' good wills marks a principled limit to any Kantian psychology. While situationists can point in the Milgram experiments to the fact that 68% of subjects shocked their partners to apparent death, no empirical study will ever be able to show that, say, 85% of those in a moral community as compared with only 10% of those outside one have good wills. In the case of the politeness hypothesis, it would be relatively easy to establish empirical measures of politeness. Psychologists could identify some set of benchmark behaviors that constitute "polite society" in a given context and observe the extent to which subjects manifest (or increasingly manifest) the benchmark behaviors.[14] However, one could not so easily establish empirical measures of virtue. Even if there were an inverse correlation between, say, violence or cheating and uses of "please" and "thank you," this finding would only be between one sort of behavior and another; we could not directly *measure* whether or not the decreased frequency of violence and dishonest is due to the development of real virtue or merely an expanded range for the polite "mere appearance of virtue." Good wills, unlike honest or beneficent behavior, can't be empirically verified.

While this is a real limitation for Kantian empirical psychology, however, Kantians should not exaggerate its scope. Empirical psychologists today make widespread appeal to mental states, personality traits, and dispositional tendencies (both conscious and unconscious). Psychology today is far from the skeptical behaviorism of the early twentieth century, and the general willingness to speculate about dispositional bases of overt behavior makes clearer than it may have seemed sixty years ago how Kant's emphasis on motive and principle could be empirically tractable. Kant, too, allows for considerable empirical knowledge of human psychological states.[15] His discussions of character are premised on the notion that character can, at least in many cases, be inferred from the consistency between a person's behavior and principles to which she is explicitly committed (e.g. 7:285). Insofar as character is a necessary precondition of a fully good will, testing subjects for consistency and stability of character indirectly tests them for the good will. We cannot know that one with a character has a good will, but we can know that one without a character lacks one, and even that the cultivation of character contributes to the cultivation of the good will (see Frierson 2006). Moreover, while Kant is skeptical about ever knowing one's *virtue*, he is not skeptical about knowing whether one is morally *evil*. Certain behaviors, patterns of behavior, and/or empirically-inferable motivational states provide empirical confirmation of human evil (see Frierson 2003:100-108, 2013:72-77). And testing human behavior can provide informative *evidence* of moral virtue, even if this evidence is indirect and imperfect. If subjects in the Milgram experiment who refuse to shock their partners share some features, this provides evidence that those features correlates with (moral) character. If it turns out that cultivating certain forms of politeness gives rise to

decreases in violence and cheating, that provides at least some empirical evidence of an improved disposition. Moreover, as hypotheses get more specific, empirical psychology can provide more specific evidence. In the case of politeness, for example, Kant claims that the way polite behavior cultivates virtue is by counteracting a tendency to excuse vice through claims that one simply could not help it. Behaving politely "betrays a self-mastery and is the beginning of conquering oneself [and thereby] a step towards virtue" (25:930). In contemporary parlance, we might say that the socially reinforced effort to become polite helps cultivate a "growth mindset" (Dweck 2006) or a sense of personal responsibility (Kohlberg and Candee 1984). Insofar as there are already tests, albeit imperfect ones, for these characteristics, and evidence that they help prevent certain sorts of situational effects (Dweck 2006, Good et. al. 2003), showing that politeness cultivates these traits provides indirect evidence that it promotes character, and thereby indirect evidence that it promotes virtue. In principle, no empirical studies could ever definitively prove that such-and-such an empirical influence causes the development of moral character. But empirical investigation can nonetheless support claims in moral anthropology, even if it never provides "complete certainty" (4:407).

Moreover, Kant's claims about inscrutability that imply the impossibility of an empirical psychology that could provide complete certainty about means for advancing what is most important in human life are not merely arbitrary declarations; they draw attention to a real limit of empirical psychology. What really matters most in life – not only moral character, but also lasting happiness, authentic relationships, and so on – is *not* empirically and systematically verifiable. Any theory that classifies moral excellence in such a way that it can be read off of behavior without the possibility of systematic error risks missing what moral life is all about. We can't take integration into an empirical-psychological model as an assumed necessary condition for a moral ideal. Whether the realization of ultimate goods is perfectly verifiable is *at least* an open question, and, for Kant, it's a question rightly answered in the negative.

These two features of any Kantian psychological research program – its philosophically controversial starting points and the empirical inscrutability of its central concepts – undermine the possibility for this research program ever to yield the wholly "fact-based" or "data-driven" program for character cultivation that is so often sought by governments, educational institutions, grant-making foundations, scientific and practical journals, and increasingly even the general public. Core elements of Kant's moral theory – his conception of the good will, the importance of character as stable commitment to principles, his eschewal of non-reflective habits and unconscious motivation, and so on – have no "data" behind them. They cannot be shown to be necessary in order to achieve some universally appreciated ideal of happiness or the good life. And while Kant sees these elements as emerging from "common moral cognition," he does not require empirical confirmation of the commonality of this cognition. Even granted these ideals, moreover, there can be no certainty about the correlation between any given empirical fact and the good will that is the ultimate goal of moralization. Thus the practical significance of any psychological data will depend upon practical sense, good judgment, and ultimately a kind of "faith." Moreover, a well-conducted Kantian empirical psychology could and should resist the widespread tendency to water-down or corrupt ideals in order to make them empirically-tractable. What matters most may not ultimately be susceptible of measurement; the best we can do is approximation and indirect cultivation. But we can, and should, use the best tools of contemporary psychology to engage in this admittedly imperfect science of (corporate) self-cultivation.

Conclusion

Contemporary psychologists may decide not to study possible mechanisms for cultivating moral character in Kant's specific sense. But the failure to pursue this research program is not dictated by anything that has been revealed in situationist (or any other) psychological research, and it is not mandated by the limitations discussed in the previous section. What situationist research shows is that most people in our society lack what Kant would call character. For philosophers like Doris and Harman, this fact implies that this sort of character is not worth pursuing. But this purported inference is a non sequitur, as evidenced by the fact that Kant, as much as any situationists today, emphasized precisely the same fact of human situation-dependence, but used it as an argument for the prevalence of human evil and the vital need for a robust moral anthropology in order to combat our self-induced fragility and impurity of will. For Kant, simply accepting our situation-dependence as an inevitable fact of human life is the worst sort of slide from frailty to depravity, from a regretted tendency to morally slip up to a complaisant acceptance that we should only be expected to be moral when it is relatively easy (in a broad and somewhat artificial sense of "easy"). The alternative is, as positive psychologists emphasize, a reorientation of empirical resources away from making excuses for behavior and towards improving character.

Almost as important as the emphasis on (inner) moral improvement rather than temptation-reduction, Kant's approach to psychology shifts away from seeing human beings as mere objects to be manipulated into certain behaviors and towards seeing them as real *people* who need to make choices for themselves. Kant's resistance to situationist arguments against character and his recognition of the ultimate limits of empirical inquiry are grounded in a direct respect for human agency. Many who appropriate situationist psychological research shift the locus of responsibility for action away from individuals and onto broader social forces. From the standpoint of psychologists and other social scientists seeking to maximize their own impact, figuring out the simplest, cheapest, most reliable, and most generalizable mechanisms for manipulating human behavior makes a lot of sense, so it is unsurprising that much psychological research is oriented in precisely this way. But Kant's philosophical orientation, at least from the time he began work on his Critical philosophy and his moral anthropology, was motivated by his conversion from one who "despised the rabble who knows nothing" to one who "would feel by far less useful than the common laborer if I did not believe that this consideration [of philosophical questions] could impart a value to all others in order to establish the rights of humanity" (20:44). Thus any Kantian psychological research program will, first and foremost, respect human agency in practice and focus research on methods for cultivation of excellent human *agents*.

Bibliography

Alfano, M. (2013) *Character as Moral Fiction*. Cambridge University Press.

Aronson, J. (2013) "Mindsets and Stereotype Threat," webinar to the *National Association for Gifted Children*, slides available at http://media01.commpartners.com/NAGC/may_2013/130514%20Presentation%20Slides.pdf

Casey, W. M. and Burton R. V. (1982) "Training Children to Be Consistently Honest through Verbal Self-Instructions" *Child Development* 53:911-19.

Cohen, A. (2009) *Kant and the Human Sciences: Biology, Anthropology, and History*, Houndmills, UK: Palgrave Macmillan .

Deci, E., Vallerand R., Pelletier, L., and Ryan, R. "Motivation and Education: The Self-Determination Perspective," *Educational Psychologist* 26:325-346.

Doris, J. (2002) *Lack of Character: Personality and Moral Behavior*. Cambridge University Press.

Duckworth, A. (2016) *Grit: The power of passion and perseverance*, New York: Simon and Schuster.

Dweck, C. (2000) *Self-Theories: Their role in motivation, personality, and development*. New York: Psychology Press.

Dweck, C. (2006) *Mindset: The new psychology of success*. New York: Ballantine Books.

Frierson, P. (under-review) "Character in Kant's Moral Psychology: Responding to the Situationist Challenge."

Frierson, P. (2014) *Kant's Empirical Psychology*, Cambridge University Press.

Frierson, P. (2013) *Kant's Questions: What is the Human Being?* Abingdon: Routledge.

Frierson, P. (2010) "Kantian Moral Pessimism" in *Kant's Anatomy of Evil*, ed. Sharon Anderson-Gold and Pablo Muchnik, Cambridge University Press, pp. 33-56.

Frierson, P. (2009) "Kant on Mental Disorder 2: Philosophical Implications," *History of Psychiatry* 20: 290-310.

Frierson, P. (2008) "Empirical Psychology, Common Sense, and Kant's Empirical Markers for Moral Responsibility," *Studies in the History and Philosophy of Science* 39:473-82.

Frierson, P. (2006) "Character and Evil in Kant's Moral Anthropology," *Journal of the History of Philosophy* 44: 623-34.

Frierson, P. (2005a) "Kant's Empirical Account of Human Action," *Philosopher's Imprint* 5.7:1-32

Frierson, P. (2005b) "The Moral Importance of Politeness in Kant's Anthropology," *Kantian Review* 9: 105-27

Frierson, P. (2003) *Freedom and Anthropology in Kant's Moral Philosophy*. Cambridge University Press.

Good, C., Aronson, J. and Inzlicht, M. (2003) "Improving adolescents' standardized test performance: An intervention to reduce the effects of stereotype threat," *Applied Developmental Psychology* 24:645-662.

Greene, J. (2008) "The Secret Joke of Kant's Soul," in *Moral Psychology, Vol. 3: The Neuroscience of Morality: Emotion, Brain Disorders, and Development*, ed. Walter Sinnott-Armstrong. Cambridge, Mass.: MIT Press, pp. 35-79.

Guyer, P. (2012) "Examples of Moral Possibility," in *Kant and Education: Interpretations and Commentary*, ed. Roth, K. and Surprenant, C., Abingdon: Routledge, pp. 124-138.

Habermas, J. (1993) *Justification and Application*, trans. C.P. Cronin, Cambridge, MA: MIT Press.

Harman, G. (2000) "Moral Philosophy Meets Social Psychology: Virtue Ethics and the Fundamental Attribution Error" in *Explaining Value and other essays in moral philosophy*, Oxford University Press.

Herman, B. (1993) *The Practice of Moral Judgment*. Harvard University Press.

Kleinberger, A. (1982) "Theories of moral action," *Journal of Moral Education* 11:147-58.

Kleingeld, P. (2014) "Debunking Confabulation: Emotions and the Significance of Empirical Psychology for Kantian Ethics" in *Kant on Emotions and Value* (ed. A. Cohen) Basingstoke:

Palgrave MacMillan.

Kleingeld, P. (2015) "Consistent Egoists and situation managers: two problems for situationism," *Philosophical Explorations* 18:344-61.

Kohlberg, L. (1981) *The Philosophy of Moral Development*. San Francisco: Harper and Row.

Kohlberg, L. and Candee, D. (1984) "The Relationship of Moral Judgment to Moral Action," in *Morality, Moral Behavior, and Moral Development* (ed. W. M. Kurtines and J. L. Gewirtz) New York: Wiley-Interscience.

Louden, R. (2011) *Kant's Human Being: Essays on his theory of human nature*, Oxford University Press.

Lopez, S. J. and Snyder, C. R. (2009) *The Oxford Handbook of Positive Psychology*. Oxford University Press.

Mazar, N., O. Amir, and D. Ariely (2008) "The Dishonesty of Honest People: A theory of self-concept maintenance," *Journal of Marketing Research* 45:633-44.

Milgram, S. (1963) "Behavioral Study of Obedience," *Journal of Abnormal and Social Psychology* 67:371-8.

Milgram, S. (1974) *Obedience to Authority: An experimental view*. New York: Harper and Row.

Miller, C. (2013) *Moral Character: An empirical theory*. Oxford University Press.

Miller, C. (2014) *Character and Moral Psychology.* Oxford University Press.

Moran, K. (2012) *Community and Progress in Kant's Moral Philosophy.* Washington, D.C.: The Catholic University of America Press.

Munzel, G. F. (1999) *Kant's Conception of Moral Character: The "Critical" Link of Morality, Anthropology, and Reflective Judgment.* Chicago, The University of Chicago Press.

Munzel, G. F. (2012) *Kant's Conception of Pedagogy: Toward Education for Freedom*. Northwestern University Press.

Peterson, C. (2006) *A Primer in Positive Psychology*, Oxford University Press.

Roth, K. and Surprenant, C. (2012) *Kant and Education: Intepretations and Commentary*. Abingdon: Routledge.

Seligman, M. and Csikszentmihalyi, M. (2000) "Positive Psychology: An Introduction" in *American Psychologist* 55:5-15.

Seligman, M., Steen, T., Park, N., Peterson, C. (2005) "Positive Psychology Progress: Empirical Validation of Interventions" *American Psychologist* 60:410-421.

Seligman, M. (2002) *Authentic Happiness,* New York: Free Press .

Shu, L., F. Gino, and M. Bazerman (2011) "Dishonest Deed, Clear Conscience: When cheating leads to moral disengagement and motivated forgetting," *Personality and Social Psychology Bulletin* 37:330-49.

Steele, C. and Aronson, J. (1995) "Stereotype Threat and Intellectual Test Performance of African-Americans," *Journal of Personality and Social Psychology* 68:797-811.

Sturm, T. (2009) *Kant und die Wissenschaften vom Menschen*, Münster: Mentis Verlag.

Wood, A. (1999) *Kant's Ethical Thought*. Cambridge: Cambridge University Press.

---

[1] This and the preceding three paragraphs summarize arguments made in xxx-references deleted for review.

[2] See Kohlberg and Candee 1984. For Milgram's original experiments, see Milgram 1963 and 1964. In essence, Milgram's experiments showed that most people are willing to torture other people for trivial reasons if instructed to do so by someone in a lab coat. Kohlberg and Candee found that whether someone is susceptible to this situational influence strongly depends on their sense of personal responsibility and their moral maturity.

[3] The classic study establishing stereotype threat effects is Steele and Aronson 1995. Steele and Aronson showed that asking test-takers about race or gender *before* a standardized exam generated a significant disparity in performance based on stereotype (e.g. black students or women doing more poorly on math than white male

students), while asking them about race and gender *after* the exam did not show such effects. This study thus confirms a very strong situation-based effect on performance. More recent studies, however, have shown that the effect of stereotype threat can be reduced, eliminated, or even reversed if students are equipped with certain mindsets, such as that "the ability in question is thought to be *improvable*" rather than "*fixed*" (Aronson 2013, cf. Dweck 2006:75-7; Good et. al. 2003). The situation-dependence of behavior here depends upon mindset characteristics that can be cultivated or combatted.

[4] For explanations of the distinction between frail, impure, and depraved wills, see Frierson 2013:78-9 or Wood 1999:285-6.

[5] For further discussion of various influences on character development, see Frierson 2006, 2014; Moran 2012; Munzel 1999. For more on Kant's pedagogy, see Louden 2011:136-49; Munzel 2012; Roth and Surprenant 2012.

[6] This would not preclude all use of examples, but it would change the *way* examples are used. Cf. Guyer 2012.

[7] In laying out these principles, I emphasize *moral* psychology, but similar points would hold for other areas of empirical research as well. Thus, for instance, one could restate (1) as the claim that "normative theory precedes psychological research" and thus use it to think about cultivating epistemic improvement or even ideals of a happy life.

[8] Empirical psychology helps specify duties that follow from the moral law for beings like ourselves. Thus, in the case of politeness, Kant goes to pains to draw on certain features of human psychology (notably our ability to be influenced by appearances that we know to be false) to carefully distinguish between illusion and deception in order to show that presenting oneself as better than one really is (polite behavior) is morally permissible because not, strictly speaking, deception (see Frierson 2005b). More generally, insofar as we have duties to promote the perfection of our faculties and the happiness of others (see e.g. 4:423-4; 6:387), we need to know what those faculties are, how to improve them, and what is conducive to others' happiness. We can know such facts only by empirical psychology. Moreover, the empirical fact that human beings generally lack character has implications for the content of our moral obligations. In particular, given that human beings are radically evil, we cannot *simply* resolve to always do what is morally correct, but must *actively struggle* against our own evil tendencies: one "remains . . . exposed to the assaults of the evil principle; and, to assert his freedom, which is constantly under attack, he must henceforth remain forever armed for battle" (6:93, cf. 6:409). This active struggle involves an additional set of obligations to morally strengthen one's will, obligations that would not be incumbent upon angels or even morally perfect human beings. As empirical psychology discovers new helps or hindrances to moral character, our general obligation to combat our own evil tendencies will give rise to further, empirically-informed obligations. Politeness, for example, is not a direct moral obligation, but it becomes morally obligatory for beings like us who have a propensity to evil but can be strengthened against that propensity through polite behavior. One final implication of humans' radical evil comes in the realm of moral appraisal, where Kant shifts from a perfectionist standard to a model wherein what it means to have a "good" will is not perfect stability but *progress*. When one considers what one ought to *do*, one should hold oneself to the highest moral standards, including the requirement that one act from a stable character. But when one looks back on one's life, or on the life of others, one can *assess* moral status in terms of a lower, or at least different, standard, one according to which progress towards perfection is a sufficient basis for moral hope (see 6:67).

[9] For Kant, of course, there is a principled problem with the whole idea that one can empirically influence the *moral* person, since one's moral personhood is ascribable to a transcendentally free empirical character. For discussion of this issue, see Frierson 2003.

[10] Even when Kant rejects physiological explanations from his pragmatic anthropology, however, he does so because we "cannot put them to use for [our] purposes" (7:119). It is less clear what would be his attitude towards physiological explanations that *could* effectively be put to use, though concerns of autonomy might lead him to prefer non-physiological interventions where possible. Cf. Frierson 2009, Frierson 2014, Sturm 2009.

[11] For discussion of this movement from various perspectives, see the articles collected in volume 55 (2000) of the *American Psychologist*, as well as Seligman et. al. 2005 and Seligman 2002. For the purpose of this paper, I focus only on Kant in relation to the primary proponents of positive psychology today, particularly Seligman. Situating a broadly Kantian positive psychology in relation to other psychologists with similar projects, including Angela Duckworth, Carol Dweck, Richard Ryan and Edward Deci, and others, would reward further study. The present paper provides only a start at this project. In addition, I have not discussed perhaps the most obvious recent psychologist for any Kantian to discuss: Lawrence Kohlberg. Kant's moral anthropology shares with Kohlberg a conception of the moral ideal as requiring a strong sense of personal responsibility and a characterization of moral problems in terms of autonomously-given moral principles. But for Kant, unlike Kohlberg, the moral ideal cannot be discerned *from* experimental research but must *guide* that experimental research, and relatedly, the primary

purpose of empirical psychology should be studying mechanisms for *improving* human beings, not merely for characterizing them. Seligman, Duckworth, Dweck, Ryan and Deci, although less specifically Kantian in the content of their moral ideals, provide better models than Kohlberg for the sort of research into character *cultivation* that Kantian moral anthropology requires.

[12] Regarding the centrality of character for moral development, see Frierson 2006, Munzel 1999, and Rumsey 1989. The importance of character for overall well-being in Kant's anthropology has not been as widely discussed. For one example of Kant's emphasis on character for well-being overall, see his *Lectures on Anthropology*, 25:1388, where he says with respect to those who lack character,

> if one does not keep his word to himself the first time, nothing is achieved. For instance, if he has resolved to get up early always and this resolution is always neglected while carrying it out, nothing is achieved. The human being who cannot believe in himself with respect to his own resolutions feels the hopelessness of not being able to procure all the good that he could have procured.

For Kant, character is the ability to govern oneself consistently by maxims, whether moral or merely prudential (e.g. 25:1387). One who lacks this ability not only fails morally, but also fails to promote his or her own well-being effectively.

[13] What is worse, if programs of cultivation are predicated on contingent and contested moral norms, the universality of those norms may end up being a self-fulfilling illusion.

[14] Experimental groups could even be divided into a control group (with no added social encouragement to be "polite"), a "mere politeness" group that learns lots of non-moral norms of politeness ("please" and "thank you," tucking in one's shirt, not wearing hats in doors, proper use of knife and fork, etc), and an "polite virtue" group that learns moral norms *as* norms of politeness. Researchers could investigate whether the effects of politeness depend upon the *sort* of politeness involved.

[15] See Frierson 2014.