

12-15-2017

The Role of the "Subject's Power" in Kant's Account of Desire

Leonard Feldblyum

Follow this and additional works at: https://scholarworks.gsu.edu/philosophy_theses

Recommended Citation

Feldblyum, Leonard, "The Role of the "Subject's Power" in Kant's Account of Desire." Thesis, Georgia State University, 2017.
https://scholarworks.gsu.edu/philosophy_theses/219

This Thesis is brought to you for free and open access by the Department of Philosophy at ScholarWorks @ Georgia State University. It has been accepted for inclusion in Philosophy Theses by an authorized administrator of ScholarWorks @ Georgia State University. For more information, please contact scholarworks@gsu.edu.

THE ROLE OF THE “SUBJECT’S POWER” IN KANT’S ACCOUNT OF DESIRE

by

LEONARD FELDBLYUM

Under the Direction of Eric Entrican Wilson, PhD.

ABSTRACT

Understanding Kant’s account of desire is vital to the project of evaluating his views about moral psychology, as well as his account of freedom *qua* autonomy. In *Anthropology from a Pragmatic Point of View*, Kant claims that “*Desire (appetitus)* is the self-determination of a subject's power through the representation of something in the future as an effect of this representation” (7:251). My goal is to clarify which of the subject’s *specific* capacities Kant means by the “subject's power,” and what role this capacity plays in desire. I argue that the subject's power cannot be her capacity to act. Rather, the subject's power is best understood as her capacity to generate the psychological states that cause action. I call these motivational states 'activation signals'. Desire consists in the self-determination of the subject’s capacity to generate activation signals by her representation of the object of desire together with an accompanying incentive.

INDEX WORDS: Immanuel Kant, German philosophy, Moral psychology, Philosophy of action, Patrick Frierson, Moral philosophy

THE ROLE OF THE “SUBJECT’S POWER” IN KANT’S ACCOUNT OF DESIRE

by

LEONARD FELDBLYUM

A Thesis Submitted in Partial Fulfillment of the Requirements for the Degree of

Master of Arts

in the College of Arts and Sciences

Georgia State University

2017

Copyright by
Leonard Feldblyum
2017

THE ROLE OF THE “SUBJECT’S POWER” IN KANT’S ACCOUNT OF DESIRE

by

LEONARD FELDBLYUM

Committee Chair: Eric Entrican Wilson

Committee: Jessica Berry

Electronic Version Approved:

Office of Graduate Studies

College of Arts and Sciences

Georgia State University

December 2017

TABLE OF CONTENTS

1	INTRODUCTION.....	1
2	KANT’S ACCOUNT OF DESIRE.....	4
3	MY ACCOUNT OF THE SUBJECT’S POWER.....	13
4	ARE WISHES DESIRES?	25
5	CONCLUSION	37
	REFERENCES.....	42

1 INTRODUCTION

Kant was the systematic philosopher *par excellence*; as is well known, his various positions in epistemology, metaphysics, ethics, and moral psychology are fundamentally interconnected. However, it has been less widely appreciated that Kant's account of empirical psychology is just as important to his philosophical system, especially to his accounts of moral motivation and freedom *qua* autonomy. In particular, one cannot understand Kant's account of how the *a priori* moral law can gain traction on our actual psychological structure, motivate us to act, and thereby ground our autonomy, without first understanding his general account of motivation and desire. So, whether one wants to vindicate Kant's positions on free will and moral motivation or refute them, one cannot do so adequately without coming to grips with his empirical psychology.

Kant's empirical psychology is his account of the psychological capacities which make it possible for human beings to live, experience, and act the way they do. His method is to figure out, mostly from the armchair,¹ the capacities human beings must have to be the way they are, and what these capacities must be like. Kant identifies and explains these capacities in terms of the *formal* role each plays in our psychology, and in terms of how they all relate to each other and work together to make human life, experience, and action possible.²

In this paper, I will contribute to our understanding of Kant's empirical psychology by clarifying an undertreated aspect of Kant's account of desire. In *Anthropology from a Pragmatic Point of View*, Kant describes 'desire' as "the self-determination of a subject's power through the representation of something in the future as an effect of this representation" (7:251). Yet he does

¹ Kant also employs his own experiences in the world, as well as the testimony of others about their experiences, including accounts of foreign peoples and their practices given in travel books (although Kant read widely, he never travelled far from Königsberg himself). See Kant's Preface to his *Anthropology from a Pragmatic Point of View* (7:120) for Kant's own account of the way he used travel books to develop his views on empirical psychology.

² For more on this, see Sections 2 and 3 below, especially footnote 21.

not explain which of the subject's various powers he has in mind. The aim of this paper is to clarify which *specific* capacity Kant refers to using the term 'subject's power'. What is at stake is how to best understand a crucial component of Kant's empirical account of the faculty of desire. In turn, understanding the faculty of desire is vital to the project of evaluating the plausibility of Kant's *moral* psychology, especially his account of moral motivation, as well as his account of freedom *qua* autonomy. So, for Kantians, I hope this paper will help clarify an important, yet neglected aspect of Kant's empirical psychology, which is crucial for understanding his accounts of moral motivation and autonomy. For non-Kantians, I hope that this paper will render an important part of Kant's empirical psychology less arcane and abstract, and make the various philosophical claims Kant bases upon it easier to understand and evaluate.

To clarify what the subject's power is, I pursue two main goals throughout this paper. The first is to block an interpretation that seems tempting, namely, understanding the subject's power as her capacity to *act*. Kant's discussions of the faculty of desire in works other than the *Anthropology*, especially his definition of the faculty of desire in the *Metaphysics of Morals*, seem to support such an interpretation. However, I'll argue that this interpretation is untenable because it clashes with other aspects of Kant's account of desire, especially his claim that wishes are desires.

My second goal is to develop an interpretation of the subject's power that avoids this problem. I'll contend that the subject's power Kant refers to in the *Anthropology* is best understood as a psychological capacity internal to the subject that is *distinct from* and *causally prior* to her power to act. Specifically, it is her capacity for generating the psychological state that motivates her to act; I call this motivational state an 'activation signal'.

The paper is organized as follows. In Section 2, I provide a basic overview of Kant's account of desire, focusing on the role that the subject's power plays in it. I also introduce my basic argument against taking the subject's power as her capacity to act: such an interpretation clashes with Kant's claim that wishes are desires. In Section 3, I develop my alternative interpretation of the subject's power, and show how it avoids clashing with Kant's claims about wishing. Finally, in Section 4 I defend my interpretation against the potential objection that wishes do not actually count as desires for Kant. For if wishes are not desires, then taking the subject's power as her capacity to act does not clash with Kant's account of wishes after all, and there is no need for my alternative interpretation. I therefore begin Section 4 by focusing on an important recent interpretation of Kant's empirical psychology offered by Patrick Frierson, who explains Kant's account of the faculty of desire in a way that rules out wishes as desires. I then argue that Frierson's attempts to explain away textual evidence that Kant does treat wishes as desires are implausible. Finally, I show that my account offers a more detailed and more plausible account of the faculty of desire.

Throughout this paper, I proceed on the assumption that when Kant refers to "a subject's power," he is referring to a *specific* capacity of hers, rather than making a vague reference to her capacities in general. I do so because Kant specifies that the subject's power is determined by certain kinds of representations (i.e., representations of the future) and not others, and that this determination must count as self-determination. Yet not all of the subject's capacities are determined by representations of the future in a way that counts as self-determination; her capacity to smell is one example of a capacity that fails to meet these specifications. So, I take it that Kant's mention of the subject's power in the *Anthropology* cannot merely be a vague

reference to her capacities in general, since this would be *too* general to meet Kant's own specifications for the role played by the subject's power in desire.

In addition, before we begin it's important to acknowledge a number of desiderata for any explanation of what Kant means by the "subject's power." First, the subject's power must be explained in terms of a capacity that is, or at least *can be*, determined by a representation of some future state of affairs (i.e., the object of desire). Moreover, this determination relation between the subject's representation of the object of desire and the subject's power must be explained such that it counts as *self*-determination. Finally, the subject's power must be explained as a *causal* power.³ With these desiderata in mind, let's begin our investigation of the subject's power.

2 KANT'S ACCOUNT OF DESIRE

In order to clarify what the subject's power is, we must first understand Kant's account of desire in general. I'll begin with four preliminary points about Kant's account of desire. The first is that Kant's empirical psychology is a *faculty psychology*. He takes as his starting point that human beings have three psychological faculties (*Vermögen*): *cognition*, *feeling*, and *desire*. Each of these faculties consists in various sub-faculties or 'basic powers'. As Patrick Frierson points out, "Kant's empirical psychology thus consists in laying out the various basic powers of soul, grouped according to his three faculties, and describing the causal laws and predispositional bases for each power" (Frierson 54). Kant then appeals to these basic powers in order to explain causal relationships between various mental states, as well as between mental states and

³ Desires often *cause* actions, so they are part of the causal order. Moreover, desires are introspectively experienced in time via inner sense, so they must be taken as part of the empirical causal order. And insofar as desires are instances of a subject's power being determined, that power must also be part of the causal order.

behavior.⁴ Although some powers can be completely explained in terms of, or reduced to, other powers, Kant claims that desire, like cognition and feeling, is an *irreducible* faculty because it cannot be completely explained in terms of the two other faculties (29:877; see also 8:181fn). So, for Kant the *faculty* of desire is a set of distinct psychological capacities, or *basic powers*, which interact in a way that generates instances of desire, and which cannot be reduced to the basic powers comprising the faculties of cognition and feeling.⁵

The second preliminary point concerns the role played by the faculty of *cognition* in producing the representation of the object of desire, which in turn determines the subject's power. Kant specifies that this crucial role is played by *foresight*, a sub-type of imagination (one of the basic powers of cognition). He states that “the faculty of visualizing something as taking place in the future is called the *faculty of foresight*” (7:183), and that “every desire contains a (doubtful or certain) foresight of what is possible through it” (7:186). In other words, the basic power of foresight is responsible for producing a visual representation of the object of desire.⁶ However, not all representations of objects of desire are *directly* contributed by foresight. This is

⁴ As Frierson puts it, “Kant aims in his psychology to *use* his catalog of human predispositions/powers to explain causal connections between mental states (and even between mental states and actions) without needing efficient-causal explanations of those predispositions' origins” (Frierson 52). For more on Kant's use of the concepts 'power,' 'capacity,' and 'faculty,' see Frierson 52-54. For a basic characterization of this method by Kant himself, see e.g. *KpV* 5:10.

⁵ In this paper, I follow Frierson in using the term ‘basic power’ to refer to the sub-faculties which constitute the three fundamental faculties of cognition, feeling, and desire. Richard McCarty argues against this use of the term ‘power’ in his 2009 (see especially pgs. 17-21). However, nothing in the arguments I present in this paper hinges on this terminological question, so I put it aside for present purposes.

⁶ Of course, the *content* of this representation might be vague or even completely indeterminate, as in the case of a peevish wish (7:251). A peevish wish is simply a restless desire for *something* (i.e., you feel the pain of a *lack*), without desiring anything in particular. However, even in the case of a peevish wish, foresight is involved. For in peevish wishing, *something* must be represented; Kant is clear that peevish wishing involves the representation of an indeterminate object of desire, rather than simply *not representing any object of desire*. It seems to me that a representation of an indeterminate object of desire just is the representation of a future feeling of satisfaction (i.e., the cessation of the feeling of restlessness characteristic of the peevish wish). If you do not represent at least that, then you have not represented an object of desire at all. So, foresight plays a crucial role in desire, because the subject must represent *something* in the future in order to have a desire at all.

because the understanding can turn the visual representations of foresight into *ideas*, which can also function as representations of objects of desire.⁷

The third preliminary point about Kant's account of desire is that there is a crucial distinction between a faculty and the *exercise* of that faculty. We have just seen that a psychological faculty is a set of basic psychological powers. In contrast, an *exercise* of a psychological faculty is a specific *instance* in which some or all of the basic powers constituting that faculty cause a mental state or behavior. This might consist in causing a specific mental state or behavior to arise in the first place. For example, the pink elephant you have just imagined is a representation that is the result of an *exercise* of your basic power of imagination, and thus of your faculty of cognition. However, the exercise of a faculty might also consist in causing a change to occur in a mental state or behavior that is already present in the subject. For example, the change in colors from pink to green in the elephant you are imagining is also an exercise of your basic power of imagination, and thus of your faculty of cognition.

The distinction between a faculty and its exercise becomes important when we turn to Kant's various characterizations of desire. Sometimes, he clearly refers to the *faculty* of desire (*Begehrungsvermögen*). For example, in the *Metaphysics of Morals* Kant says: “the *faculty of desire* is the faculty to be, by means of one's representations, the cause of the objects of these representations” (6:211).⁸ In contrast, it seems as though Kant's discussion of “desire” in the

⁷ In his discussion of foresight, Kant claims that “empirical foresight is the *anticipation of similar cases (expectatio casuum similium)* and requires no rational knowledge of causes and effects, but only the remembering of observed events as they commonly follow one another, and repeated experiences produce an aptitude for it” (7:186). This suggests that there is also another kind of foresight which *does* involve rational knowledge of causes and effects, and which therefore involves not only imagination but also the understanding. Specifically, the subject recognizes a pattern in the observed events recalled through imagination, and forms an idea about this pattern, in which the pattern is understood in terms of cause and effect. It seems to me that this process of abstraction is what *rational foresight* would involve, which explains how foresight can give rise to *ideas* that can function as non-visual representations of objects of desire.

⁸ Kant offers almost identical characterizations of the faculty of desire in *KpV* (5:9fn), and in the *Lectures on Metaphysics* (29:1012).

Anthropology, in which he mentions the subject's power, concerns the *exercise* of the faculty of desire, i.e., the production of an *instance* of desire (*die Begierde*). Since Kant does not specify that he means the *faculty*, his account of desire in the *Anthropology* seems to be a characterization of what an *instance* of desire is like.⁹

Nonetheless, I take Kant's remarks about instances of desire in the *Anthropology* to be a source of information about the *faculty* of desire that is just as useful as Kant's explicit remarks about the faculty itself in other works. After all, the faculty of desire is the subject's *general capacity* to produce instances of desire with the features that Kant describes in the *Anthropology*. From this, it follows that the faculty of desire must include whatever sub-faculties (i) are needed to produce phenomena with those features, and (ii) are not part of cognition or feeling. For example, although the basic power of foresight is needed to produce instances of desire, it is part of the faculty of cognition, which means that it cannot be part of the faculty of desire. In contrast, Kant specifies that the subject's power necessarily plays a role in producing an instance of desire. However, the subject's power is not a sub-faculty of cognition or feeling.¹⁰ So, we can conclude that the subject's power is part of the faculty of desire. In this way, the fact that the subject's power is part of the faculty of desire can be gleaned from Kant's characterization of an instance of desire in the *Anthropology*.

The final preliminary point concerns the role of the faculty of *feeling* in desire. Once the subject represents an object of desire, that representation causes her to feel pleasure or pain. Kant

⁹ This claim is supported by the fact that immediately after the characterization of desire, Kant describes different kinds of desires, including *inclinations*, *wishes*, *idle wishes*, and *peevish wishes*. This seems more like a taxonomy of different ways the faculty of desire can be *exercised* than a description of the faculty itself.

¹⁰ This is suggested by the fact that Kant opens Part III of the *Anthropological Didactic*, his discussion of the *faculty* of desire, with a description of the exercise of that faculty as involving determination of the subject's power. Moreover, the fact that he opens his discussion of the faculty of desire with a description of the exercise of that faculty suggests that he sees information about a faculty's exercise as indicating something about the faculty itself as well.

calls this feeling of pleasure or pain the “impelling cause” of her desire (28:254).¹¹ Based on the *source* of the representation which gives rise to the impelling cause of a desire, Kant then distinguishes between the *higher* and *lower* faculties of desire (28:254). As Frierson helpfully explains, “insofar as a desire is the direct result of the senses or imagination, it is part of the 'lower' faculty of desire. Insofar as it proceeds from the understanding or reason, a desire falls under the 'higher' faculty” (Frierson 62).

However, to fully understand the role of feeling in Kant’s account of desire, we must say more about the *impulses* that are the impelling causes of desire. Kant calls these impelling causes of desires *incentives* (29:895).¹² As Patrick Frierson explains, “these impelling causes operate by means of feelings of satisfaction and dissatisfaction” (Frierson 2014, 62).¹³ Yet for Kant, feelings of satisfaction and dissatisfaction just are feelings of pleasure and pain, respectively.¹⁴ Moreover, Kant claims that the feelings of pleasure and pain just are the impulses to maintain one's present state or leave one's present state, respectively (7:231). So, an incentive is a feeling, specifically an impulse to either maintain or leave one’s present state (i.e., pleasure or pain, respectively), that is felt by the subject in inner sense, and which causes her to form a desire.

¹¹ We’ll see below that Kant calls these impelling causes ‘incentives’. For Kant, incentives are the impelling causes of desires both because they are a necessary (though not sufficient) cause of desires arising in the first place, and because it is in virtue of having an incentive that you are impelled (i.e., caused) to pursue your desire by acting (unless your desire is a wish).

¹² Kant also characterizes incentives as “the subjective ground of a desire” (4:427), and contrasts incentives to *motives*, which are “the objective ground of volition” (4:427). A motive is objective because it is not a mental state; rather, it is the object, state of affairs, or fact which constitutes a *reason* for you to desire. And a volition is an instance of *willing*, and the will is the faculty of desire governed by the understanding, which insofar as it is pure is called *reason* (5:55). That is why motives are the objective ground of *volition*, rather than desire more generally.

¹³ See also Kant's claim that “*Concupiscence* (lusting after something) must also be distinguished from desire itself, as a stimulus to determining desire. Concupiscence is always a sensible modification of the mind but one that has not yet become an act of the faculty of desire” (6:213). Concupiscence is a psychological state in which one *feels pain* (lust *qua* feeling of lack and dissatisfaction), such as hunger. Insofar as this feeling of pain comes to serve as the subjective ground and impelling cause of a desire, it is an incentive. So 'concupiscence' and 'incentive' are two different terms for the same psychological state; the former simply refers to the state itself, while the latter emphasizes its relationship to other psychological states.

¹⁴ See Kant's account of pleasure and pain in §60 of the *Anthropology* (7:231), as well as his discussions of moral satisfaction in §64 (7:237) and of aesthetic satisfaction in §69 (7:244), both of which are phrased in terms of pleasure and pain.

Now, in the same paragraph of *Anthropology* §73 in which he describes an instance of desire, Kant also names and describes a variety of types of desire. One of these is the *peevish wish*: “The undetermined desire, in respect of the object (*appetitus vaga*), which only impels the subject to leave his present state without knowing what state he then wants to enter, can be called the *peevish wish* (one that nothing satisfies)” (7:251). Taken together with Kant's account of pain as the urge to leave one's current state (7:230-231), Kant's account of peevish wishes suggests the following picture of how the faculty of desire works. The faculty of desire consists in the subject's potential for changing states by having her 'power' determined by certain interactions between her faculties of feeling and cognition. Specifically, this interaction is between a feeling of *pleasure or pain*, and a cognition of an *object of desire* (which, in the case of a peevish wish, is indeterminate). The upshot is that for Kant, all desires consist in the subject's power being determined by representations of future states of affairs together with incentives. Likewise, the faculty of desire consists in the subject's *capacity* for her subject's power to be determined by this sort of interaction between cognition and feeling.

So far, we can summarize our preliminary account of the faculty of desire as follows. The faculty of desire is a set of basic powers, which cannot be completely reduced to the basic powers comprising the faculties of cognition and feeling. These basic powers work together to produce instances of the psychological state called 'desire'. Specifically, certain kinds of representations of future states of affairs (i.e., representations of objects of desire) work together with feelings of pleasure and pain (i.e., incentives) to causally determine a basic power called the “subject's power” to cause actions which can bring about the subject's object of desire in the world. Moreover, Kant holds that this entire process counts as self-determination.

If the faculty of desire is understood in this way, an interpretation of the subject's power and its role in the faculty of desire seems to follow straightforwardly. In order to realize one's desires, i.e., bring about the objects of one's desires in the world, one must *act*. The sort of action through which one can realize one's desires is interaction with the world via bodily movement. Thus, in *any* instance of desire, the subject's representation of an object of desire can only cause her to bring about that object in the world if it causes her to interact with the world through bodily movement. Since Kant states in the *Anthropology* that the subject's representation of the object of desire specifically determines the subject's power, it seems natural to take the subject's power as her power to interact with the world through bodily movement, i.e., her capacity to act. In turn, it seems at first glance that the faculty of desire is the subject's capacity to cause herself to act, and thereby to bring about certain states of affairs in the world, by representing those states beforehand.

One advantage of this interpretation of the subject's power is that it easily accounts for Kant's definition of the faculty of desire in the *Metaphysics of Morals*. However, it clashes with Kant's account of wishing in the *Anthropology*. In that work, Kant states that wishing is “*desiring without exercising power to produce the object*” (7:251).¹⁵ Since the power to actually produce the object is the power to interact with the world through bodily movement, Kant's claim here is that wishes are desires in which the power to move, i.e., the power to act, is not exercised. But we have also seen that in the *Anthropology*, Kant claims that desire always involves the self-determination of the subject's power. So, if wishes are desires, then the subject's power cannot

¹⁵ In the *Metaphysics of Morals*, Kant claims that “the faculty of desire in accordance with concepts, insofar as the ground determining it to action lies within itself and not in its object, is called a faculty to *do or to refrain from doing as one pleases*. Insofar as it is joined with one's consciousness of the ability to bring about its object by one's action it is called *choice*; if it is not joined with this consciousness its act is called a *wish*” (6:213). Given the extensive qualifications preceding Kant's characterization of wishing in this passage, especially the fact that he is talking not about the faculty of desire, but only the faculty of desire “in accordance with concepts, insofar as the ground determining it to action lies within itself and not in its object,” I take the characterization of wishing in the *Anthropology* as more informative and useful for present purposes.

be the power that actually produces the object, i.e., her capacity to interact with the world via bodily movement. For in order to produce the object of her desire in the world, the subject must exercise her power to act – that is, she must actually move. Yet wishing is a kind of *self-determination* of the subject's power (i.e., a desire) in which she does not actually *exercise* her power to move. In some cases, she might not even be *able* to exercise her power to move in a way that would bring about the object of her desire; Kant calls these *idle wishes* (*ibid.*). For example, you might desire to stop reading this paper right now, even though you don't exercise your ability to do so; this would be a *wish*. Or, you might desire to draw a triangle with four sides; this would be an *idle wish*. According to Kant, both of these are desires, yet neither causes you to exercise your capacity to bring about what you desire by moving your body. The upshot here is that if wishes are desires, then *desires do not necessarily bring about the actuality of the object of desire* by causing action. This means that the subject's power, which plays a role in *all* instances of desire, cannot be her capacity to act.¹⁶

In sum, Kant's account of desire in the *Anthropology* seems to suggest two things. First, the subject's power cannot be the power to act. Otherwise, Kant's account of desire in the *Anthropology*, which counts wishes as desires, would be incoherent. Second, there must be a distinction between *self-determination* of the subject's power, which according to the *Anthropology* is necessarily involved in any instance of desire (including wishes), and the exercise of the subject's physical power to act (by using her body, moving her muscles, etc.), which *is not* necessarily involved in desire.

¹⁶ Another reason to reject the notion that the subject's power is the capacity to act, is that the notion of a 'capacity to act' is vacuous. At best, it is shorthand for a host of specific psychological and physical capacities, which *includes* the faculty of desire, and even more specifically, the subject's power that Kant refers to in the *Anthropology*. This is the case even if we explain the capacity to act in terms of the capacity for bodily movement, which, while less vacuous than 'capacity to act', is still just shorthand for many specific capacities.

However, these two points seem to clash with Kant's claim in the *Metaphysics of Morals* that "the *faculty of desire* is the faculty to be, by means of one's representations, the cause of the objects of these representations." For that claim makes it seem as though the subject's power just is the capacity for bodily movement which brings about the objects of her desires. So, our investigation of the subject's power faces a dilemma. On one hand, if we follow the *Metaphysics of Morals* and interpret the subject's power as her capacity to act, then it seems that we cannot account for Kant's claim in the *Anthropology* that wishes are desires, and must instead explain it away. On the other hand, if we follow the *Anthropology* by taking wishes as desires, then we cannot take the subject's power to be her capacity to act, and thereby bring about the objects of her desires in the world, which Kant emphasizes in the *Metaphysics of Morals*.

The *Metaphysics of Morals* was published contemporaneously with the *Anthropology*; it is unlikely that in two books published in the same year, Kant both does and does not count wishes as desires, and offers incompatible accounts of the faculty of desire. So for the moment, let's see whether there is a plausible way to interpret the subject's power as a power other than her capacity to act, in a way which does not clash with Kant's explicit claim in the *Anthropology* that wishes are desires. After investigating this possibility in Section 3, in Section 4 we will compare the resulting account to Patrick Frierson's interpretation of the faculty of desire. According to Frierson, Kant thinks that all desires cause actions which bring about the object of desire in the world (in the absence of hindrances), and that wishes aren't desires. By comparing my interpretation from Section 3 with Frierson's, we will be able to determine which is most plausible, and thus resolve our interpretive dilemma.

3 MY ACCOUNT OF THE SUBJECT'S POWER

In this section, I attempt to identify the specific capacity that Kant refers to in the *Anthropology* as the “subject’s power.” Since my aim is to do so in a way which does not clash with Kant's claim that wishes are desires, I begin by assuming *that* the subject's power is something other than her capacity to act. Next, I examine Kant’s notion of self-determination in order to shed light upon *how* the subject’s power is distinct from her capacity to act. I then propose that the subject’s power is best understood as her capacity for generating the psychological state that pushes her to perform a particular action; I call this motivational state an *activation signal*. I conclude the section by explaining this claim and elaborating upon it, in order to develop a detailed account of the subject’s power and its role in the faculty of desire.

In order to develop our account of the subject's power, let's assume for the moment that the subject's power cannot be her capacity to act.¹⁷ Under such an assumption, we must treat the determination of the subject's power, and the determination of her capacity to act, as determinations of two distinct powers. So what sort of determination is the determination of the subject's power, and how is it different from determination of her capacity to act? The fact that Kant calls determination of the subject's power *self-determination* provides a first clue. Desire is something the subject does to herself before she acts in the world, if she even acts at all. Self-determination is not the sort of relation between subject and world that is part and parcel of the power to act. So, the self-determination at work in desire is a relation *internal to the subject*. In other words, desire is a step *causally prior to action* (or inaction), and involves determination of the subject’s power by something else within her.

¹⁷ Of course, the capacity to act is still involved in many instances of desire, namely those desires that, unlike wishes, cause bodily movements which bring about the object of desire. Below, I call such desires ‘action-causing desires’.

Moreover, recall that desire plays a role in Kant's *empirical* psychology of action. This means that the sense of 'self-determination' at work in desire cannot be the sense in which *autonomy of the will* is self-determination. The will's autonomy is its capacity for a *certain kind* of self-determination, namely *self-legislation*, i.e., being determined by representations of the moral law (4:440). But the moral law is *a priori*, so the self-determination at work in desire, which is empirical, cannot always be the type of self-determination which autonomy is the capacity for (namely self-legislation).¹⁸ Rather, the sense of self-determination we are interested in is broader, because it encompasses what happens in *most* cases of desire, namely, all desires that are *not* caused by feelings of respect for the moral law. So, when Kant says that desire involves self-determination, he means that desire involves *causal* determination of certain psychological processes of the human organism *by other such processes* (as opposed to non-psychological bodily processes or events external to the body). That is, the sort of self-determination relevant to our discussion of the subject's power is *determination by psychological processes internal to the subject*.¹⁹ For example, if you imagine a horrible monster and this causes you to feel fear, the causal relationship between your representation of the monster and your feeling of fear counts as self-determination in the relevant sense, but *not* as self-determination *qua* autonomy of the will.

¹⁸ In fact, according to Kant, self-legislation is at work in desire *only* when the subject feels respect as a result of representing the moral law; the subject's *feeling* of respect and *cognition* of the moral law then cause the subject to desire to act in accordance with duty (see 5:78-81).

¹⁹ I have in mind here the sense of 'internal' that Kant refers to with the word 'within' when he talks about "a *comparative* concept of freedom (according to which that is sometimes called a free effect, the determining natural ground of which lies *within* the acting being)" (5:96). See also Kant's use of 'internally' when saying that the faculty of desire is "where life manifests itself not merely in feeling, internally..." (7:286); since feelings are known through inner sense, I take there to be a meaningful distinction between *internal* and *external* phenomena, where internal phenomena are known via inner sense, and external phenomena are known via the five traditional senses. In making this claim, I take no stance as to the metaphysical status of these processes. By Kant's lights, claims about these processes, and thus claims about psychological self-determination, are claims about *phenomena*.

Now, for the self-determination relation at work in desire to be internal to the subject's psychology, its relata must be internal to her psychology as well. One of these relata – the representation of the object of desire – is clearly internal. For the representations which determine the *lower* faculty of desire – sensations and imagined “representation[s] of something in the future” – are known via inner sense. The same is true of the representations which determine the *higher* faculty of desire, such as concepts and judgments. The subject's own faculty of cognition produces these representations, and the subject accesses them via inner sense. Finally, the *incentives* accompanying these representations are also internal, because they are feelings of pleasure or pain that the subject accesses via inner sense.

As for the other relatum – the subject's power – we have assumed for the time being that it cannot be her capacity to act. This means that the subject's power must be a power internal to the subject that is distinct from, and prior to, her capacity to act. I propose that this power is best understood as the subject's capacity for generating the psychological state that *pushes* her to perform a particular action. In other words, the “subject's power” is her capacity to generate an *activation signal* which, barring interference from other factors, causes a specific performance of the bodily movements called 'action'. Activation signals just are motivations to perform *specific* actions, and the subject's power is her capacity to generate these motivational states.²⁰

²⁰ In contemporary neuroscience, such activation signals are chains of various kinds of signals, including *action potentials*. An action potential is an electrical chemical impulse that travels down a single (presynaptic) neuron's axon, and causes synaptic firing which may influence another (postsynaptic) cell, whether by exciting it or inhibiting it. A chain of such signals from neuron to neuron can terminate at a particular region of the body (e.g. the arm) and result in an action, behavior, reflex, or other bodily process. So *one way* to take my claim that the subject's power consists in her capacity for generating activation signals, is that the subject's power is the capacity of certain neurons in her brain to set into motion chains of action potentials in response to other parts of her brain representing an object of desire and signaling pleasure or pain, with the terminal action potentials causing certain muscles in her body to contract, thereby generating certain kinds of bodily movement. However, exploring this connection to contemporary neuroscience is beyond the scope of this paper; for more, see e.g. Frijda 2007 (especially pgs. 4-7, where Frijda identifies “the principle of passion: to manifest states of action readiness”), Ridderinkhof (forthcoming - especially pgs. 5-7), and Blakemore & Vuilleumier (forthcoming). In any case, given that our topic is Kant's *empirical* psychology, it seems useful and important to draw out these sorts of connections to relevant scientific research.

I contend that Kant has such a role in mind for the subject's power when he characterizes the will (i.e., the faculty of desire in accordance with concepts) as “a faculty either of producing objects corresponding to representations or of determining itself to effect such objects (whether the physical power is sufficient or not), that is, of determining its causality” (*KpV* 5:15). If taken in isolation, the first disjunct might lead us to understand the subject's power as her capacity to act (i.e., her capacity to actually produce the object of her desire). However, the second disjunct makes clear that the subject can form a desire even if she is unable to act in a way that produces the object of her desire (whether unable in general, or unable due to some interfering factor present in that specific instance of desire). The second disjunct also suggests that the subject's capacity to determine herself to bring about the objects of her desires (and thereby ‘determine her causality’) is a distinct capacity that is required for, and must be exercised prior to, her capacity to act. By taking the subject's power as her capacity to generate an activation signal, which then goes on to determine her capacity to act (barring interfering factors), we can account for the distinct capacity for self-determination of one’s causality that Kant thinks is required for, and causally prior to, one’s “physical power” to act.

The *specificity* of the actions caused by activations signals – that is, the fact that a given activation signal causes Action A and not Action B - is a result of the subject's representation of a particular object of desire. As we saw above, the subject’s representation of an object of desire is her representation of some future state of affairs through her basic power of foresight. My claim is that the *content* of this representation is primarily responsible for the specificity of an activation signal. So for instance, if the future situation Alice represents is one that she isn’t used to, or one that is very important for her, then Alice’s representation of that situation prior to her generation of an activation signal might include a visualization of specific actions she’ll take.

This visualization, or some other similar represented content (e.g., a plan 'verbally' expressed in Alice's internal monologue), causes her 'subject's power' to generate an activation signal to *specific* muscles in her body to make *specific* movements.²¹ Yet in other cases, the represented situation is one that Alice is used to, and to which she has developed habitual ways of responding as a result of repeated exposure to similar situations. In such cases, Alice is already habituated to act immediately, and in very specific ways, without first needing to represent an elaborate plan of action. That is what it means to have a *habitual desire*, which Kant calls an *inclination* (7:251).

Consider the following example. On a hot day you hear the music of an ice cream truck, and immediately begin walking towards the sound. In this case, you respond immediately because you happen to have an inclination to eat ice cream on hot days that you developed over numerous prior exposures to such situations. Part of what it means to have an inclination is to have the tendency to immediately generate activation signals in response to certain kinds of stimuli, even when those stimuli cause you to form only vague, simple representations of the future. So, you don't need to represent an elaborate plan of action in order to respond immediately to the ice cream truck's music in a specific way. In any case, whether the desire in question is stand-alone or the result of an inclination, the subject's representation of the object of desire is a crucial part of her desire, because it contributes specificity to her activation signal

²¹ Note that, since Kant's psychology of action is a faculty psychology, it is enough for present purposes to specify the capacities which generate the psychological phenomena under investigation, as long as the resulting account is compatible with whatever physiological knowledge about the brain turns out to be relevant. It is beyond the scope of this paper to explain *how* the content of a representation results in the specificity of an activation signal; as long as it turns out *that* this causal link exists, the plausibility of the account provided here is not threatened. In general, on my view Kant does not need to commit himself to any particular account of the physiological mechanics of how activation signals are produced and delivered to the muscles. All Kant needs to claim is that, since he is providing a psychology of *action*, at some point the psychological phenomena he discusses must link up with the body and physiological explanations of it, since the point is to explain how psychological states can cause movement (and, through movement, the object of desire). That is, these claims about activation signals are formal; as long as there is some physiological set of mechanisms that carries out the functions that Kant specified, his account is not problematic merely because it does not commit to a particular *physiological* account of activation signals.

such that it can be the psychological mechanism that actively pushes her to perform a *specific* action. In the case of inclinations, over time a specific activation signal becomes habitually (i.e., physiologically) associated directly with the subject's reception and recognition of relevant stimuli, such that when the subject represents those stimuli, that specific activation signal is immediately generated.

In Section 2, we discussed Kant's notion of an incentive, and the role incentives play in his account of desire. At first glance, the difference between incentives (which Kant calls "impelling causes" of desire) and what I call activation signals might not be obvious. However, activation signals are distinct from incentives in the following way. Incentives are feelings of pleasure or pain that merely push you to maintain or leave your present state (respectively); they do not by themselves cause *specific* actions. In contrast, activation signals are the psychological mechanisms that motivate you to perform *specific* actions (rather than the more general *leaving or maintaining your present state*). Thus, activation signals are not merely incentives, and the capacity to generate activation signals is distinct from the mere capacity to feel incentives (which is part of the faculty of feeling). However, although activation signals are not merely incentives, generating activation signals does involve incentives. Together with representations of objects of desire, incentives cause the subject's basic power of generating activation signals (her 'subject's power') to *actually generate* activation signals. In other words, incentives and representations of objects of desire are causally prior to activation signals in Kant's psychology of action.²²

This last claim about the role of incentives and representations of objects of desire in causing the generation of activation signals can be understood in two distinct ways, which generate two different versions of my account. In the first version, we can take the subject's

²² Likewise, the faculties of feeling and cognition are causally prior to the faculty of desire in general (and the subject's power in particular) in Kant's psychology of action.

power to generate activation signals as a *discrete psychological mechanism* that is distinct from the mechanisms which produce representations of objects of desire and incentives, and that is determined by their interaction. Since in this version of my account, the subject's power acts as a middleman between cognition and feeling, on one hand, and activation signals on the other hand, I call this first version the 'Middleman interpretation'. On the Middleman interpretation, activation signals are produced by the subject's power when it is triggered to do so by certain interactions between her faculties of cognition and feeling. A useful analogy would be the way that a television's receiver (the mechanism which receives signals from the remote) is distinct from the mechanisms (my finger and the remote) whose interaction triggers the receiver to prompt the television screen to display images. In this scenario, the receiver's prompt to the television screen is analogous to the activation signal from the subject's power to her muscles.

In the second version of my view, we can take the subject's power to generate activation signals merely as her *capacity to undergo a certain kind of interaction* between her representations of objects of desire and her incentives. When this interaction takes place, it directly generates an activation signal which (in the absence of interfering factors) determines the subject's body to move, which is why I call this second version of my account the 'Direct interpretation'. On the Direct interpretation, the subject's power is not a discrete psychological mechanism that acts as a middleman. Rather, it is simply the subject's capacity to undergo the kind of interaction between cognition and feeling that directly produces activation signals. In turn, on the Direct interpretation, an activation signal can be characterized as *the way in which* an incentive and a representation of an object of desire work together to directly cause a specific exercise of the subject's capacity for movement (in the absence of interfering factors). Here, a useful analogy would be the way that one's capacity to light a pile of wood on fire just is one's

capacity to initiate a certain kind of interaction between the flint and steel in one's tinderbox, namely the spark-producing kind of interaction. This interaction directly generates a spark which, absent interfering factors such as strong wind or rain, is enough to light the pile of wood on fire. In this scenario, the spark is analogous to the activation signal.

It seems to me that to settle whether the Middleman interpretation or the Direct interpretation is correct, we would first need to identify the precise neural correlates of the subject's power in the brain. Then, we would need to see whether those correlates constitute a psychologically or functionally distinct module,²³ or an emergent capacity for directly producing activation signals through certain interactions between cognition and feeling (which does not require a physiologically distinct middleman). Fortunately, for present purposes there is no need to settle which interpretation is correct. For on either version of my view, three key claims are made. First, what causes an exercise of one's capacity for bodily movement is an activation signal. Second, what generates the activation signal is the subject's power (understood as distinct from one's capacity for bodily movement). Third, the process as a whole through which the activation signal is generated and attempts to cause movement is made up of a set of relations that are internal to the subject.²⁴ So, in both versions of my view, my core distinction between the *determination of the capacity to generate activation signals*, and the further *determination of the capacity to act*, is maintained and explained.

The fact that this distinction is maintained on both versions of my view is important, because it allows us to explain Kant's claim that a wish is “*desiring without exercising power to*

²³ For more about mental modules and how they can play a role in a faculty psychology, see Fodor (1983), especially pgs. 2-38.

²⁴ This process is experienced by the subject in inner sense as a discrete mental state, namely as an instance of desire (I'll say more about this below).

produce the object” (7:251).²⁵ On my view, when you desire, your representation of the object of desire, together with the accompanying incentive, causes the generation of an activation signal to your muscles. Sometimes this signal reaches its destination and your body moves. In that case, your physical power to act was exercised to produce the object of your desire, and your desire was an *action-causing* desire.²⁶ Other times, the signal fails to reach its destination, in which case your physical power to act wasn’t exercised to produce the object, and your desire was a wish.

We can explain an activation signal’s failure to reach its destination by appealing to an interfering psychological or physiological factor. There are many ways this can work; I will provide two examples illustrating how interfering factors can cause one’s desire to be a wish from the start, as well as how interfering factors can turn action-causing desires into wishes.²⁷

First, suppose that Alex imagines winning a million dollars from a lottery ticket, and feels pain (i.e., feels an incentive) due to his lack of the lottery ticket. Alex’s representation of a winning ticket, together with his feeling of pain, together cause the generation of an activation signal to Alex’s muscles to begin moving in a way that will lead him to acquire a lottery ticket. However, at that moment Alex also represents to himself the extremely low probability of winning the lottery, as well as the high cost of lottery tickets, the inconvenience of having to leave his comfy chair to drive to the store, etc.. All of these representations cause him to feel a new pain, which interferes with the activation signal and prevents it from determining Alex’s

²⁵ Notably, Kant does not specify that the “power” which is not exercised in wishes is the subject’s power; I take him to be referring here to the “physical power” to act (5:15).

²⁶ At least in §73 of the *Anthropology*, Kant lacks a dedicated term for a desire in which the subject’s power to produce the object is actually exercised (as opposed to ‘wish’, which is his term for a desire in which the subject’s power to produce the object is *not* exercised). So, I will use the term ‘action-causing desire’ to refer to this specific type of desire. (Contrast this terminological move to Frierson’s account, which takes ‘desire’ to just *mean* ‘action-causing desire’ for Kant (Frierson 56). As we will see below, this causes problems for his interpretation.)

²⁷ See also Kant’s claims that “*fright* is suddenly aroused fear that *disconcerts* the mind” (7:255, emphasis added), and that *surprise* “at first impedes the natural play of thought” (7:261). I take it that *one* of the various ways fright or surprise can interfere with mental processes in the manner Kant describes is by interfering with activation signals, and thereby preventing them from reaching the muscles.

muscles to move in a way that will lead him to purchase the ticket. At most, Alex experiences a slight impulse to get out of his chair, which his body does not follow. At this point, Alex certainly *wants* a winning lottery ticket. However, the sense in which he wants one is that he *wishes* he would acquire one. The upshot is that, since no action was caused by that wish, at no point did Alex have an *action-causing desire* to acquire a winning lottery ticket. This example illustrates how one's thoughts can cause a desire to be a wish from the start, and prevent that wish from developing into an action-causing desire.

In our second example, Izzy has a strong desire for the use of pesticides in agriculture to be banned. She takes actions such as volunteering for anti-pesticide groups in order to bring about her object of desire. At this point, she clearly has an action-causing desire for pesticides to be banned. But eventually, she becomes disillusioned with the methods of anti-pesticide groups, and comes to believe that there are no effective or legitimate methods currently available for bringing about the banning of pesticides. This new belief interferes with Izzy's actually doing anything more to bring about the banning of pesticides, and she ceases all anti-pesticide activity. However, she never stops *wanting* pesticides to be banned, because she continues *wishing* that they would be banned. Prior to her disillusionment, Izzy had an action-causing desire to ban pesticides. After her disillusionment, this desire transformed into a wish, because there was now an interfering factor: her representation of the lack of effective and legitimate methods. This example shows how interfering factors can turn action-causing desires into wishes.

Collectively, these examples should show that my account of the subject's power straightforwardly explains how wishes count as desires, in a way that avoids the problems that arise if we take the subject's power as her capacity to act. My account does not treat all desires as

action-causing, does not rule out wishes as a kind of desire, and most importantly for present purposes, it does not identify the subject's power with the capacity to act.

Nonetheless, my account remains compatible with all of Kant's characterizations of desire which we examined above, including those that motivate the view that the subject's power is her capacity to act. In particular, my account explains Kant's definition of the faculty of desire in the *Metaphysics of Morals* as “the faculty to be, by means of one's representations, the cause of the objects of these representations” in the following way. One's representations of objects of desire generate activation signals. These activation signals then cause the muscles of one's body to perform specific movements that cause the actuality of the objects of one's desires in the world. Without these activation signals, there is no movement, and hence no realization of the objects of one's desires. That is why the faculty of desire is one's capacity to be the cause of one's objects of desire being realized in the world. Yet not all activation signals successfully determine the body to move; when they fail to do so due to some interfering factor, the desire in question is a wish. So, my account is compatible with Kant's remarks about desire in the *Anthropology* as well.

Finally, note that on my view, the subject's power is *merely* her capacity to generate activation signals. In contrast, the faculty of desire is the subject's capacity for carrying out the entire process I have described. That is, the faculty of desire is the subject's capacity to generate activation signals *through* representing some future state of affairs as painful or pleasant to her (i.e., together with an incentive). This means that, on my account, an instance of desire is an instance of a representation by the subject of some future state of affairs as painful or pleasant to her causing her to generate an activation signal. An activation signal is not an instance of desire; it is merely the output of the basic power that Kant calls the subject's power. The subject's power

is determined to generate an activation signal by certain kinds of interaction between her faculties of cognition and feeling. These psychological processes happen very fast, and are not accessible to introspection; we can't detect activation signals in inner sense. Instead, what we experience in inner sense is a *desire*; that is, we experience an *indistinct* representation of the process as a whole, in which the composition of the representation remains unclear (7:137-138). It is this indistinct representation of the process as a whole that comprises the singular experience we call 'desire'. In any case, an activation signal cannot be the same thing as an instance of desire. Kant clearly specifies that an instance of desire includes a representation of an object of desire. But activation signals are not such representations, nor do activation signals contain such representations. So, an activation signal is merely a *component* of a desire that is not itself discernible through inner sense. Only the instance of desire as a whole is detectable in introspection. Likewise, the basic power to generate activation signals - the subject's power - is distinct from the faculty of desire as a whole, and is merely a component of that faculty.

In sum, I contend that the "subject's power" is a causal power internal to the subject that is distinct from and causally prior to her capacity to act. I suggest that this power is best understood as the subject's capacity to generate motivational states which I call 'activation signals'. This capacity is causally determined by the interaction of the subject's faculties of cognition and feeling into *attempting to further cause* an exercise of her body's capacity for interaction with its environment through movement, by sending an activation signal to her muscles. The faculty of desire, then, is the subject's capacity to generate activation signals by representing some future state of affairs as painful or pleasant to her (i.e., she represents the object of her desire together with an incentive).²⁸ In any instance of desire, the subject's

²⁸ According to Kant, it does not matter whether the representation of the object is causally prior to the incentive, or *vice versa*; both relationships are possible (6:211-212).

generation of an activation signal is *self*-determination (in the broad sense), because it is caused by a representation of an object of desire that is produced by her own faculty of cognition, and by a felt incentive produced by her own faculty of feeling, both of which are internal to the subject's psychology.

4 ARE WISHES DESIRES?

A potential objection to my interpretation of the subject's power is that, if Kant does not actually count wishes as desires, then we don't need to reject the notion that the subject's power is simply the capacity to act. In that case, we wouldn't need to talk about activation signals at all in order to understand the subject's power and its role in desire. One reason to think that wishes aren't desires is that treating them as desires fails to capture the strong connection between desire and action suggested by Kant's definition of desire in the *Metaphysics of Morals*. There, Kant defines the faculty of desire as "the faculty to be, by means of one's representations, the cause of the objects of these representations."²⁹ At first glance, it seems to follow from this definition that *any* exercise of the faculty of desire – that is, any instance of desire – must actually cause an action which brings about the object of desire in the world. In that case, it seems that wishes, which *do not* cause desire-realizing actions, cannot count as desires in Kant's sense. Now recall that on the view I developed in Section 3, a desire may or may not cause a desire-realizing action, depending on whether or not the subject's activation signal was blocked from reaching her muscles by an interfering factor. If her desire does cause an action, it is an *action-causing desire*; if her desire fails to cause an action, it is a *wish*. So, my account of the subject's power seems to clash with Kant's definition of the faculty of desire in the *Metaphysics of Morals*

²⁹ See also Kant's claim that "all desires have a relation to activity and are the causality thereof" (25:1514).

because it treats wishes as desires, despite the fact that wishes do not cause the subject to perform desire-realizing actions.

In fact, many contemporary commentators interpret Kant's account of the faculty of desire in a way that rules out wishes as desires.³⁰ It seems to me that this position has the status of a received view in the secondary literature, mostly because the topic of wishes remains undertreated.³¹ So, in this section, I want to confront it, not only in order to defend my view against a potential objection, but also because of the crucial implications of this issue for Kant's moral psychology and account of freedom. We have to get Kant's empirical psychology right in order to get his other views right too.

Perhaps the most thorough interpretation of the faculty of desire in the secondary literature according to the received view is given by Patrick Frierson in his monograph *Kant's Empirical Psychology*. I focus on it because Frierson is one of the few commentators who has backed his ruling out of wishes as desires with an extensive discussion of Kant's views about empirical psychology in general. Frierson's account of the faculty of desire aims to capture the strong connection between desire and action suggested by many of Kant's remarks about that faculty, most importantly his definition of it in the *Metaphysics of Morals*. For Frierson, that definition suggests that any *instance* of desire must actually cause a desire-realizing action. If so, then wishes, which *do not* cause such actions, cannot count as desires in Kant's sense.

Frierson thus makes two key claims about the faculty of desire. The first is that (1) according to Kant, a psychological state counts as a desire only if it causes an action which

³⁰ See, e.g. Stephen Engstrom (2009: 27), who claims that Kant defines desire as "*efficacious representation*" which brings about the actuality of its objects. If Engstrom is right, then Kant's definition of desire presupposes that desire always causes action (and thus that the subject's power is the capacity to act), since it is through action that the actuality of objects of desire is brought about. In that case, if Engstrom is right, then Kant rules out wishes as desires.

³¹ However, Richard McCarty is one example of a commentator who *does* treat wishes as desires (see McCarty 2009: 15-17).

brings about a represented state of affairs in the world. In turn, this means that (2) there is an important sense in which *wishes aren't desires*. As Frierson puts it:

desires simply *are* action-oriented representations that serve as '*cause of the actuality of the object*' of representation (29:1012), so Kant's notion of desire is more closely connected to choice and action than the customary English sense of desire, whereby one can desire something without actually pursuing it. Once one has a desire in this general Kantian sense, one is committed to action, and action follows necessarily in the absence of hindrances. One might, for example, desire a mango and then find oneself unable to climb the tree, but one's representation will not count as desire unless it prompts one to action. In contrast to typical English usage, for Kant desires mark an end to deliberation, not factors taken into account in deliberation. Thus a Kantian 'desire' (*Begehren/Begierde*) might better be called a volition (as Kant does, identifying it with *Wollen* at 25:1334). Kant develops other categories – such as 'wish' (*Wunsch*), 'inclination' (*Neigung*), and 'ground of desire' – that serve the purpose of what we might call 'mere' desires. (Frierson 56)

For Frierson, any interpretation of Kant's views about desire according to which "one can desire something without actually pursuing it" conflates the broad, everyday sense of 'desire' with Kant's narrower notion of desire, which is "more closely connected to choice and action." That is why Frierson claims that wishes aren't desires in Kant's sense of the term. Since my account of the faculty of desire *does* treat wishes as desires, Frierson might say that my account fails to capture the narrow, action-causing sense of desire that he attributes to Kant. Moreover, if Frierson is right, then the subject's power that Kant refers to in the *Anthropology* might be the capacity to act after all, in which case there is no need for the account I offer above in Section 3.

Both Frierson and I share the same project: to explain Kant's psychology of action, and thereby set the stage for an explanation of Kant's *moral* psychology (see Frierson 54). Given this shared project, I acknowledge that one desideratum of any account of Kant's views concerning the faculty of desire is that it must explain the strong relationship between desire and action. However, any plausible interpretation of Kant's views must also explain all of the relevant textual evidence, or else *explain away* textual evidence that seems to undermine that

interpretation. We have seen that, in the *Anthropology*, Kant explicitly classifies wishes as desires. So, a second desideratum is that any account of Kant's views about the faculty of desire must either explain why wishes count as desires *despite* the important relationship between desire and action, or else explain away that passage in the *Anthropology*, as well as any other similar passages.

Given his overall interpretation of Kant's views about desire, Frierson must *explain away* evidence that Kant thinks wishes are desires. One of the strongest such pieces of evidence is this crucial footnote in the *Critique of the Power of Judgment*, in which Kant responds to critics of his definition of desire in the *Metaphysics of Morals* who pointed out that not all desires are action-causing:

the definition of the faculty of desire as **the faculty for being through one's representations the cause of the reality of the objects of these representations** has been criticized because mere **wishes** are also desires, but yet everyone would concede that he could not produce their object by their means alone. - This, however, proves nothing more than that there are also desires in a human being as a result of which he stands in contradiction with himself, in that he works toward the production of the object by means of his representation **alone**, from which he can however expect no success, because he is aware that his mechanical powers [...] are either inadequate or even aimed at something impossible.... (5:177-178fn)³²

Here we have a passage in which Kant explicitly claims that a wish is a specific kind of desire. More precisely, a wish is a desire that fails to bring about the object of desire due to an interfering factor: the subject's awareness of the inadequacy of his own powers to bring it about. For example, my desire for world peace is a wish because of my awareness that I am powerless to bring about world peace. My representation of my powerlessness acts as a *hindrance* which prevents my desire from causing actions which would bring about, or at least increase the likelihood of, world peace.

³² See also Kant's unpublished First Introduction to the *Critique of the Power of Judgment* (20:230fn), in Kant 2005.

We have seen above that Frierson tries to account for the role of such hindrances by claiming that “once one has a desire in this general Kantian sense, one is committed to action, and action follows necessarily *in the absence of hindrances*” (Frierson 56, emphasis added). However, this claim causes tensions to emerge in Frierson’s interpretation. For taken on its own, Frierson’s claim suggests that, in the *presence* of hindrances, action does not necessarily follow from a desire. In that case, one *can* desire something without actually pursuing it. Yet we have also seen above that *in the very same passage*, Frierson *denies* that one can desire something without actually pursuing it, and consequently rules out wishes as desires. At first glance, it is difficult to see how Frierson’s denial that one can desire something without pursuing it coheres either with the textual evidence, or with Frierson’s own interpretation of that evidence.

Frierson seems to respond to this tension in his account by claiming that, although not all desires necessarily *cause* actions, all desires necessarily *commit one to action*. In that case, one cannot desire something without actually pursuing it, as long as ‘pursuing the object of desire’ is understood as *having a commitment to act in a way that would bring about the object of desire*.³³ And if wishes do not cause one to pursue the object of one’s desire in the sense of committing one to act, then wishes cannot count as desires. Of course, the plausibility of this argument hangs on what Frierson thinks a ‘commitment to action’ is. According to Frierson:

In general, a desire is a commitment to action that is realized in the absence of external impediments or subjective incapacity. In the case of wishes, one is aware of incapacities at the moment of desire, and hence this commitment has no effect in action. Even in these cases, Kant suggests that “powers are repeatedly strained by representations in order to make their object real” (20:231n), so mere wishes are not wholly inert. When different and mutually incompatible grounds of desire are operative in a particular case, only one of these will give rise to a desire in the strict sense. But other grounds of desire will give rise to a “yearning [Sehnsucht]” (20:231n)

³³ Returning to Frierson’s own example, his claim is that one cannot desire a mango without being committed to act in a way that would lead to one’s acquiring a mango, even if one finds oneself unable to climb the tree and actually acquire a mango. So, being ‘committed to action’ in Frierson’s sense must be something other than actually performing the action in question.

that can have real and lingering effects (such as “overexciting and enfeebling the heart” (20:231n)) even if these effects are not tied to actually bringing about its object [...]. (Frierson 83)

Here, Frierson suggests that a ‘commitment to action’ is characterized by *repeated straining* of the subject’s powers by her representations of objects of desire, which pushes those powers to cause actions which bring about those objects of desire in the world. In the case of what Frierson calls a “desire in the strict sense”, i.e., an action-causing desire, this repeated straining will actually cause an action. However, in the case of wishes, this repeated straining will give rise to real and lingering physiological effects that might not have anything to do with the object of desire, such as excitement or restlessness. Kant himself characterizes wishes in terms of such physiological effects in the rest of the footnote we examined above:

Although in the case of such fantastic desires we are aware of the inadequacy of our representations (or their unsuitability) to be **causes** of their objects, nevertheless their relation as causes, hence the representation of their **causality**, is contained in every **wish**, and it is especially visible if this is an affect, namely **longing**. For the latter prove by the fact that they expand the heart and make it flaccid and thus exhaust our powers that the powers are repeatedly strained by means of representations, but the mind, in view of the impossibility, is inexorably allowed to sink back into exhaustion.... (5:177-178fn; see also 20:230fn)

Other than associating commitments to action with repeated straining of the subject’s powers, Frierson does not offer additional clarification of what a ‘commitment to action’ is. This causes problems for his interpretation. After all, if a ‘commitment to action’ consists in the straining of the subject’s powers, yet this straining amounts merely to the causation of some sort of physiological effect in the subject (whether action or affect), then given Kant’s claims about the physiological effects of wishes, wishes clearly ‘commit one to action’. In that case, wishes count as desires, even if Frierson were correct that one cannot desire something without actually pursuing it, that what it means to pursue something is to have a commitment to action, and that therefore all desires necessarily commit the subject to action.

To an extent, Frierson anticipates my line of argument by trying to explain away Kant's remarks about the physiological effects of wishing. Specifically, Frierson tries to show that the straining in wishes is different from the straining in desires, such that the physiological effects of wishes are not enough to count as commitments to action, and thus do not indicate that wishes are desires. First, he appeals to Kant's distinction between *live* and *dead* determining grounds of desires. *Live* grounds of desire are sufficient for bringing about actions, while *dead* grounds of desire are insufficient for bringing about actions, due to some impeding factor which hinders them from causing actions.³⁴ Using this distinction, Frierson says the following about wishes and their physiological effects:

the “dead” grounds for desire are not wholly inert. Instead, Kant develops a further category for thinking about pseudovolitions that do not actually impel one to action. In an unpublished introduction to his Critique of Judgment, Kant discusses “a phenomenon which is certainly noteworthy for empirical psychology,” that “there are also determinations of the faculty of desire in which it is in contradiction with itself” (20:230n). He describes this phenomenon with the concept of a “wish [Wunsch]” (7:251), a sort of “desiring without exercising power to produce the object” (7:251), which is possible only when one lacks a “consciousness of the ability to bring about one's object” (6:213; see also 20:230n; 25:206, 577–8, 795, 1109–10). (Frierson 83)³⁵

³⁴ As Frierson puts it, “In his lectures on metaphysics, [Kant] explains that grounds (powers) can be either “living” or “dead.” With dead powers, “the determining ground for an effect is internally [but not fully] sufficient” (28:565); a “dead power is a ground whereby an effect can exist but does not exist due to external circumstances” (29:824). Often these circumstances consist of a “positive impediment” where “there [is] a real power whose effect is an even stronger object, or a hindrance, something which opposes the effect of a given power” (29:825; see also 28:565). In such cases, one might have a “ground [that] is internally sufficient but the effect is still missing” (29:825)” (Frierson 82).

³⁵ Here, Frierson's interpretation of the cited passages from Kant contains a number of inaccuracies. In 6:213, Kant merely claims that “The faculty of desire in accordance with concepts, insofar as the ground determining it to action lies within itself and not in its object, is called a faculty to *do or to refrain from doing as one pleases*. Insofar as it is joined with one's consciousness of the ability to bring about its object by one's action it is called *choice*; if it is not joined with this consciousness its act is called a *wish*.” It is unclear why Frierson takes this to be an exhaustive characterization of Kant's concept of ‘wish’, given that here Kant is not talking about the faculty of desire in general, but only about the faculty of desire *in accordance with concepts*. As Frierson himself points out, “desires can be caused by a variety of types of cognition, from raw sense-perceptions [...] to principles of reason” (Frierson 84), and *all* of these kinds of desires can occur without exercising power to produce the object, not only desires “in accordance with concepts.” Moreover, this passage only supports the following claim: *if* the faculty of desire in accordance with concepts is not joined with consciousness of one's ability to bring about the object of desire through one's action, *then* that exercise of the faculty of desire in accordance with concepts is a wish. Yet Frierson claims here that wishing is possible *only* when one lacks a consciousness of the ability to bring about one's object. So, Frierson not only elides Kant's crucial qualifier “in accordance with concepts,” but he also uses Kant's words in 6:213 as evidence for a claim that those words cannot justify. He is not alone in doing so. Richard McCarty similarly

So, when taken together, Frierson's (2014: 82-83) claims yield the following argument concerning the physiological effects of wishes:

- (1) The physiological effects of wishes are not caused by the straining of *live* (i.e., unhindered) grounds of desires, but rather by the straining of *dead* (i.e., hindered) grounds of desire.
 - (2) Only the straining of *live* grounds of desire will give rise to a desire in Kant's sense of the term, since only live grounds of desire are unhindered from causing action, and for Kant all desires cause actions.
 - (3) The straining involved in wishes does not give rise to a desire in Kant's sense of the term. (FROM PREMISES 1 AND 2)
- Therefore,
- (4) The physiological effects of wishes do not indicate that wishes are desires in Kant's sense of the term.

However, this argument faces a significant obstacle: there is reason to think that Premise (2) is false, and that according to Kant, dead grounds of desire do give rise to desires, such that wishes *do* count as desires. Recall that Frierson rules out wishes as desires in the first place because he wants to account for Kant's claim, most notably in the *Metaphysics of Morals*, that the faculty of desire is "the faculty to be, by means of one's representations, the cause of the objects of these representations." Well, in his response to Friedrich Bouterwek's review of the *Metaphysics of Morals*, Kant heavily qualifies that definition in a way that weakens the support it lends to Premise (2) of Frierson's argument. Regarding what sort of causality is at work in instances of desire, Kant says:

are there not also intense but still consciously futile longings (e.g., Would to God that man were still alive!), which are *devoid of any deed* but not *devoid of any result*, since they still work powerfully within the subject himself (make him ill), though not on external things? A desire, as a *striving (nisus)* to be a *cause* by means of one's representations, is still always causality, at least within the subject, even when he sees the inadequacy of his representation for the effect he envisages. (6:356)

elides the qualifier "in accordance with concepts," for reasons that are equally mysterious (see McCarty 2009: 15). Julian Wuerth's claim that "idle" desires (i.e., *wishes* – a term which Wuerth eschews without explanation) become active "if chosen" seems to reflect this elision as well (see Wuerth 2014: 231).

Here, Kant explicitly acknowledges that although any desire is a striving to be a cause by means of one's representations, the striving in question is not always a striving to cause *action*. Rather, some desires, specifically *wishes* (of which longings are a sub-type – see 7:251), are strivings to be a cause of *results other than action*, including physiological results such as making one ill. Here, the relevant sense of striving is that of pressure, exertion, or *straining*, as indicated by Kant's use of the Latin word '*nisus*.' This sense of striving suggests merely that desire involves a brute causal push within the subject that will necessarily produce some kind of effect. In the case of wishes, the straining forces they involve “work powerfully within the subject himself,” producing effects such as illness. For example, one's unsatisfied longing for one's dead friend to come back to life can cause an effect such as depression. And crucially, this sort of causality is *enough* for a psychological state to count as a desire for Kant (as long as it includes the other factors Kant mentions throughout his works). That is, as long as it is not *devoid of result*, a psychological state can still count as a desire, even if it is *devoid of deed*.³⁶

The fact that for Kant, a psychological state can still count as a desire even if it is devoid of deed, means that Frierson's Premise (2) is false. Frierson is wrong when he claims that only live grounds can give rise to desire. Even if a psychological state is produced by a *dead* grounds

³⁶ Immediately after this vital passage, Kant says: “Since, however, all that is in question here is the relation of a cause (a representation) to an effect (a feeling) in general, the causality of a representation (whether the causality is external or internal) *with regard to its object* must unavoidably be thought in the concept of the faculty of desire” (6:357, emphasis added). At first glance, the qualification “with regard to its object” seems incompatible with the interpretation I have just advanced, according to which it is enough that a desire involves some kind of causality and generates some kind of effect. However, Kant seems to be claiming only that the effect of the representation cannot be merely a result of *having represented something*, regardless of what was represented. That is, in an instance of desire, the effect generated by the representation of the object of desire must have been a result of the *particular* representation in question; had the content of the representation been different, the effect would have been different. So, in the case of one's longing for a dead friend to return to life, the completely internal effect of the representation – one's feeling of depression - is a result of the content of the representation. In the similar case where one instead longs for a dead friend to enjoy everlasting bliss in heaven, the effect of the representation will be a different feeling, since the content of the representation is different.

of desire, it is nonetheless a desire, since it is not devoid of results such as psychological and physiological effects internal to the subject.

In sum, Kant does not claim that one cannot have a desire without actually attempting to bring about the object of desire in the world (i.e., “pursuing” the object of desire). Rather, Kant claims that once one has a desire, some result will be brought about *either* in the world *or* within the subject, as an effect of the causal striving (i.e., straining of the subject’s powers) at work in all desire. In other words, Kant’s reply to Bouterwek makes it clear that desires are not *commitments to action*, but rather brute physiological *strainings or strivings* to cause effects which can include specific actions, but are not limited to them. Given that Kant’s reply to Bouterwek is part of an attempt to clarify his definition of the faculty of desire in the *Metaphysics of Morals*, there seems to be little reason to take that definition as supporting Frierson’s claim that “desires simply *are* action-oriented representations,” nor any of the other iterations of the received view in the secondary literature. And if desires are not necessarily action-oriented representations, then there is little reason to accept Frierson’s claim that wishes are not desires, which blocks the potential objection to my account with which we began this section.

My interpretation, in contrast to Frierson’s, neatly accounts for Kant’s reply to Bouterwek. On my view, Kant’s claim that a psychological state can still count as a desire as long as it is not *devoid of result*, even if it is *devoid of deed*, is his acknowledgment that as long as an activation signal is generated by the subject’s power (through the interaction of her faculties of cognition and feeling), the resulting psychological state accessible to her through introspection counts as a desire. When the activation signal successfully reaches the subject’s muscles and causes her body to move, her psychological state is not devoid of deed, so it is an action-causing

desire. But if an interfering factor blocks the activation signal from reaching her muscles, her psychological state is devoid of deed, and is therefore a wish. However, other physiological consequences internal to the subject will be brought about by this blockage; for example, other parts of her brain can be affected in ways that can “make one ill.” In this way, on my account wishes are devoid of deed, but are not devoid of the kinds of *results* that Kant thinks make wishes count as desires.

To illustrate this point, let’s consider the case of a restless child in class. The child wishes that he could play outside, but does not move from his desk; the activation signal that would push him to move is blocked, for example by his fear of getting in trouble with the teacher. So, the child's desire is devoid of (the relevant) deed. But the blocking of the child’s activation signal to move has further physiological and psychological effects. In terms of feelings, the child might experience a steadily increasing frustration; in terms of behavior, the child might tap his fingers on the desk or start doodling. We can account for these effects using Kant’s description of desires as involving striving *qua* pressure, exertion, or straining. Whatever physiological forces are at work in the child's brain when an activation signal is generated, they do not simply disappear when that signal is prevented from reaching his muscles; instead, they find another way to discharge themselves.³⁷ So, just because the desire’s effect isn’t a desire-realizing action, but rather is some other effect internal to the subject, does not make the child’s desire something other than a desire.³⁸

The upshot of all this is that, if I am right about the implications of Kant’s reply to Bouterwek, then Frierson fails to plausibly explain away Kant’s various remarks about wishes

³⁷ Even if an activation signal is suppressed by even stronger physiological counter-forces, e.g., by sudden fright at the moment an activation signal is generated, the activation signal will have *some* physiological effect, however minute (most likely in the brain).

³⁸ This should not surprise us, given that the sort of self-determination at work in desire is, as we saw in Section 3, simply causal determination by psychological processes internal to the subject.

and their physiological effects. Thus, Frierson erroneously rules out wishes as a kind of desire, and erroneously treats all desires as action-causing. It seems to me that Frierson commits these errors because he is not sufficiently clear about the basic powers involved in the faculty of desire. We have seen that Frierson tries to account for the relationship between desire and action using his vague, and ultimately problematic, notion of ‘commitment to action’. He does this in order to *ensure* that desire be connected to action, so that the faculty of desire can play a certain role in his reconstruction of Kant’s psychology of action. However, Frierson ends up simply packing action-causation into Kant’s definition of desire - a move which is not only question-begging, but also fails to actually explain Kant’s account of desire. For the notion of a ‘commitment to act’ is vacuous; at best, it can work as shorthand for a variety of specific psychological and physiological phenomena which, as we have seen, Frierson does not adequately explain. So, although my dispute with Frierson is at least partly terminological, it is not trivial. I dispute Frierson’s use of terminology in order to correct his errors, and thereby ensure that discussion of Kant’s moral psychology proceeds from an accurate and internally consistent account of his empirical psychology of action.

Unlike Frierson’s interpretation, my account avoids making these errors because it is more precise about the basic powers that are involved in the faculty of desire. In particular, my account identifies the basic power *in virtue of which* desire is strongly related to action, namely, *the subject’s power*. My account also specifies that it is the subject’s power, understood as her specific capacity to generate activation signals, which in the case of wishes gets hindered by interfering factors, such that the activation signal does not reach her muscles. This fits neatly with Kant’s claim that a “dead power is a ground whereby an effect can exist but does not exist due to external circumstances” (29:824); in the case of wishes, the capacity to generate activation

signals is a dead power with respect to causing an effect in the outside world via the subject's muscles. Nonetheless, my account treats wishes as desires. On my account, the faculty of desire just is the subject's capacity to generate activation signals by representing an object of desire together with an incentive. So, a wish is clearly an exercise of the faculty of desire, since in any instance of wishing an activation signal is generated which simply fails to reach the subject's muscles. That is, on my account wishes count as desires, because whether or not a psychological state counts as a desire has nothing to do with whether or not the activation signal actually reaches the subject's muscles and causes her to act. However, the fact that the subject's capacity to generate activation does cause her muscles to move *when it is unhindered* (i.e., when it is a *live* power), explains the strong connection between desire and action. Thus, the greater precision of my interpretation means that it satisfies both desiderata for any interpretation of Kant's views about the faculty of desire which we mentioned at the start of this section. In turn, I hope that satisfying these desiderata lends support to my contention that the best way to understand the subject's power is as a capacity *prior* to her capacity to act, namely, the subject's capacity to generate activation signals.

5 CONCLUSION

I'll conclude our investigation of the subject's power by summarizing our results and saying more about why they matter. In this paper, I tried to accomplish two goals. My first goal was to block the tempting interpretation of the subject's power simply as her capacity to *act*. I argued that this interpretation is untenable because it clashes with other aspects of Kant's account of desire, especially his account of wishing. I then defended that argument against Patrick Frierson's claim that for Kant, wishes do not count as desires. I hope to have shown that, given

Kant's qualifications of his definition of the faculty of desire, a psychological state can count as a desire even if it fails to cause a desire-realizing action. We cannot rule out wishes as desires merely because they are devoid of deed. Instead, we must treat wishes as desires, just as Kant does in the *Anthropology*. And if wishes count as desires after all, then we cannot take the subject's power simply as her capacity to act.

If I have successfully accomplished this goal, then at least two important implications follow. The first is that showing that wishes count as desires after all has significant consequences for how best to understand Kant's account of *conflicts of desire*. Consider Frierson's claim that, according to Kant, "one cannot strictly have a desire that conflicts with another, or a desire that conflicts with an active commitment to duty" (Frierson 83). Frierson makes this claim because he thinks that desires necessarily cause actions, and in any case of conflicting *grounds* of desire only one will be live, so a subject can only have one desire (in Kant's sense) at a time.³⁹ However, if I am right that for Kant, wishes are desires after all, then the conflict between my wish to be at the beach right now, and my desire to finish writing this paper, is in fact a conflict of desire, even though only the latter desire is currently causing me to act. So, showing that wishes count as desires suggests that Kant thinks there can be genuine conflicts of desire. In turn, this affects how we should understand Kant's views about issues such as weakness of will and moral motivation. Explaining all this in detail seems like a fruitful topic for future investigation.

³⁹ See Frierson's claims that "instincts, inclinations, and character all function as different grounds of desire, and these can come into conflict. When one has a strong instinct for one thing and a weaker inclination or principled commitment to something else, one's instinct will be "live" and one's inclination or character will be present, but only as a "dead" power. Strictly speaking, one will not have desires that conflict. One cannot strictly have a desire that conflicts with another, or a desire that conflicts with an active commitment to duty, but one can have an inclination that would ground a desire but for the presence of another ground that conflicts with it. In cases where one underlying ground of desire conflicts with another, we choose (and thus desire in the fullest sense) one object but not the other" (Frierson 82-83).

The second important implication of ruling out the notion that the subject's power is simply her capacity to act, is that doing so makes it necessary to identify which specific capacity the subject's power actually is. And in fact, providing my own account of the specific capacity that Kant refers to as the 'subject's power' was my second main goal in this paper, which I pursued in Section 3. There, I argued that the subject's power is a causal power internal to the subject that is distinct from, and causally prior to, her capacity to act. I suggested that this power is best understood as the subject's capacity to generate activation signals. This capacity gets causally triggered by the interaction of cognition and feeling into *attempting to further cause* an exercise of the subject's capacity for interaction with its environment through movement, by sending an activation signal to her muscles. In turn, an instance of desire is the subject's generation of an activation signal through representing some future state of affairs as painful or pleasant to her. In any instance of desire, the subject's generation of an activation signal is *self-determination* because it is caused by a representation of an object of desire and an incentive, both of which are internal to her psychology.

Showing that the subject's power is a causal power internal to the subject's psychology that is distinct from, and causally prior to, her power to act has significant implications for how best to understand Kant's account of freedom. In *Groundwork* 4:460, Kant characterizes *two interrelated problems concerning freedom* as impossible to answer: (1) How can the moral law determine a specific feeling of pleasure/displeasure; and (2) How does this feeling (i.e., respect) generate a *motivation* to act according to duty? This does not trouble Kant in the *Groundwork*, since in Section III of that work he claims that "freedom must be presupposed as a property of the will of all rational beings" (4:447), and then explains "autonomy of the will *as the supreme principle of morality*" (4:440) in terms of "the concept of freedom" (4:446). However, in the

later *Critique of Practical Reason*, Kant begins by quickly showing that there is a Fundamental Law of Pure Practical Reason (the representing of which is a “fact of reason” (5:31)), and proceeds to explain freedom in terms of the fact that the will can be determined by a representation of this Fundamental Law. This suggests that Kant moved away from the pessimism about solving the two problems above that he displays in *Groundwork* 4:460, and towards a position that Stephen Engstrom describes as follows: “the moral law, conceived as the will's objective determining ground, is the representation of how the will ought to be determined; that same law, conceived as the subjective determining ground of a given subject's will, is the same representation *actually determining* that subject's will. So far as the moral law is a subjective determining ground of the will, it is a spring. It is a cognitive representation, but one that has force, or efficacy” (Engstrom 2010, 93). That is, the moral law has motivational force. And since freedom consists in an agent’s capacity to act according to the moral law, explaining *how* the moral law motivates us such that we can desire to act in accordance with it is also vital for explaining why Kant claims that human beings are bound for practical purposes to take themselves as free (5:57).

Crucially, explaining how the moral law gains traction on our psychology and motivates us to act in accordance with it, requires that we understand Kant's account of psychology of action in general. For we can only understand what happens in the special case when the representation determining the faculty of desire is a representation of the moral law, if we already understand how the faculty of desire works in general. Yet this means that we must first understand Kant's account of the faculty of desire and its relationship to action, which requires us to properly understand the subject's power whose determination is what Kant calls 'desire' in the *Anthropology*. If my account of the subject’s power is correct, then the following basic picture of

Kantian moral motivation emerges. The subject represents the moral law; the content of this representation is *a priori*. Once it is generated, the representation of the moral law causes a feeling of respect to arise in the subject. Together, the subject's representation of the moral law and her feeling of respect trigger an exercise of the subject's power to generate activation signals, and the subject forms a desire to act in accordance with her moral duty.

Of course, explaining this process in detail, especially how the subject can represent something *a priori* in the first place, is an immense project in its own right, and offers another promising area for future investigation. In any case, my hope is that this paper will contribute to the discussion concerning how best to understand the subject's power, such that investigation of Kant's *moral* psychology can be undertaken from a firmer understanding of Kant's *empirical* psychology. Finally, I hope that by identifying the subject's power as her capacity to generate the motivational states that cause her to perform specific actions, I have made Kant's psychology of action easier to grasp for Kantians and non-Kantians alike.

REFERENCES

1. Blakemore, R.L. & Vuilleumier, P. (forthcoming). "An emotional call to action: Integrating affective neuroscience in models of motor control." *Emotion Review*.
2. Engstrom, S. (2009). *The Form of Practical Knowledge: A Study of the Categorical Imperative*. Cambridge, Mass: Harvard UP.
3. Engstrom, S. (2010). "The *Triebfeder* of pure practical reason." In *Kant's Critique of Practical Reason: A Critical Guide*, ed. Andrews Reath and Jens Timmermann. Cambridge: Cambridge UP. 90-118.
4. Fodor, J.A. (1983) *The Modularity of Mind*. Cambridge, Mass: MIT Press.
5. Frijda, N. H. (2007). *The Laws of Emotion*. Mahwah: Erlbaum.
6. Frierson, P.R. (2014) *Kant's Empirical Psychology*. Cambridge: Cambridge UP.
7. Kant, I. (2006) [1798]. *Anthropology from a Pragmatic Point of View*, trans. Robert B. Louden. Cambridge: Cambridge UP.
8. Kant, I. (2015) [1788]. *Critique of Practical Reason*, trans. Mary Gregor. In *Kant: Practical Philosophy*. Cambridge: Cambridge UP.
9. Kant, I. (1996) [1785]. *Groundwork for the Metaphysics of Morals*, trans. Mary Gregor. In *Kant: Practical Philosophy*. Cambridge: Cambridge UP.
10. Kant, I. (1997). *Lectures on Metaphysics*, trans. Karl Ameriks and Steve Naragon. Cambridge: Cambridge UP.
11. Kant, I. (2005). *Notes and Fragments*, trans. Curtis Bowman, Paul Guyer, and Frederick Rauscher. Cambridge: Cambridge UP.
12. McCarty, R. (2009). *Kant's Theory of Action*. Oxford: Oxford UP.
13. Ridderinkhof, K.R. (forthcoming). "Emotion in Action: A Predictive Processing Perspective and Theoretical Synthesis." *Emotion Review*.
14. Wuerth, J. (2014). *Kant on Mind, Action, and Ethics*. Oxford: Oxford UP.