



The Principle of Reason's Self-Preservation in Kant's Essay on the Pantheism Controversy

ABSTRACT: *In his 1786 essay on the pantheism controversy, 'What Does It Mean to Orient Oneself in Thinking?', Kant implies that 'the maxim of reason's self-preservation [Selbsterhaltung]' is reason's first principle for orienting itself in thinking supersensible objects. But Kant does not clearly explain what the maxim or principle of reason's self-preservation is and how it fits into his larger project of critical philosophy. Nor does the secondary literature. This article reconstructs Kant's discussion of the principle of reason's self-preservation in 'What Does It Mean to Orient Oneself in Thinking?' It suggests that this principle is best understood as the discipline of pure reason. The principle of reason's self-preservation performs the same methodological function that Kant assigns to the discipline of pure reason. This principle establishes the rule of law in reason and subjects reason to its own laws. In so doing, it prevents reason's dialectical errors and also grounds reason's faith (Vernunftglaube), which in turn systematically conditions the practical use of reason.*

KEYWORDS: Kant, critical method, discipline, reason's self-preservation, faith

Introduction

In the last footnote of his 1786 essay on the pantheism controversy, 'What Does It Mean to Orient Oneself in Thinking?', Kant refers to 'the maxim of reason's self-preservation [Selbsterhaltung]' (Kant 2001: 18 [WDO 8:146]).¹ This maxim is Kant's most fundamental answer to the question he poses in the title of the essay. The maxim of reason's self-preservation is the first principle of reason to orient itself in thinking supersensible objects. Yet Kant says little about what the principle of reason's self-preservation is and how it fits into his critical philosophy. Throughout his corpus, Kant uses the term *self-preservation* often in naturalistic

I thank several anonymous reviewers, especially the two reviewers of *Journal of the American Philosophical Association*, and Rachel Bryant for their comments and criticisms that helped me improve this article. I also thank Ellen Wert for her editorial suggestions and copyediting of this article.

¹ With the exceptions of the translations from Kant's *Reflexionen zur Anthropologie* and *Reflexionen zur Logik*, which are mine, I use the recent translations from the Cambridge Edition of the Works of Immanuel Kant in consultation with the corresponding works in the German-language De Gruyter *Gesammelte Schriften* (widely known as the Akademie edition), occasionally modifying the translations. In addition to the page citations from the recent editions, for those who might use other translations or original-language editions, I provide in brackets the volume and page numbers from the Akademie edition.



and moral senses. In its naturalistic sense, *self-preservation* is used in association with inclination (see, for example, Kant 2000: 246 [KU 5:374], Kant 1999a [Refl I 15:632]). In its moral sense, *self-preservation* broadly refers to a use of practical reason to prevent moral corruption (see, for example, Kant 2000: 145 [KU 5:261], Kant 1996b: 545 [MS 6:419]). But Kant refers to *self-preservation as a principle of reason* only once, in ‘What Does It Mean to Orient Oneself in Thinking?’ If the principle of reason’s self-preservation is as significant as Kant implies, why does he discuss it only in passing in a footnote at the end of a short essay on a controversy the cause of which seems so adventitious to his philosophy? Why does not Kant mention self-preservation as a principle of reason in any other text published in his lifetime? The place and significance of the principle of reason’s self-preservation in critical philosophy remains unclear. Even interpreters who engage Kant’s 1786 essay usually do not mention, let alone examine, this principle (see, for example, Beiser 1987; Di Giovanni 1992; Zammito 1992; Munzel 1999; Deligiorgi 2005; Ferrarin 2015; Chance and Pasternack 2018). Rudolf Langthaler seems to be the only interpreter who takes seriously Kant’s reference to the principle of reason’s self-preservation. Discussing the programmatic significance of Kant’s mention of this principle in his *Nachlass* for the development of his later philosophy of religion, Langthaler argues that the principle of reason’s self-preservation is fundamental for the transition from critique to true metaphysics (Langthaler 2011, 2018). But Langthaler focuses on Kant’s philosophy of religion. He does not discuss the *methodological*, particularly the *disciplinary*, function of the principle of reason’s self-preservation.

My primary goal in this article is to show the methodological, especially the disciplinary, function of the principle of reason’s self-preservation. Kant develops his philosophy of religion and his conception of reason’s faith on the basis of reason’s self-preserving discipline. The principle of reason’s self-preservation is best understood as the discipline of pure reason, that is, the first of the methodological or ‘formal conditions of a complete system of pure reason’ (Kant 1997: 627 [KrV A708/B736]). Through self-discipline, reason establishes the rule of law in reason, and preserves and orients itself in thinking supersensible objects. In addition to clarifying Kant’s argumentation in the pantheism controversy, I offer an exposition of the disciplinary nature of the principle of reason’s self-preservation, demonstrating that the discipline of pure reason is the first principle for the systematicity of not only the theoretical use but also the practical use of reason and thus Kant’s philosophy of religion.

In ‘What Does It Mean to Orient Oneself in Thinking?’, Kant reframes the central question of the pantheism controversy—the question of the relation between faith and reason—as an internal (methodological) question of human reason. He also criticizes Friedrich Jacobi’s response to this question. Jacobi claims that the freedom to think, including freedom in acts of faith, requires freedom from the constraint of the laws of reason. In his criticism of Jacobi, Kant argues that the freedom to think consists in such constraint—that is, in the rule of law in reason. He views the rule of law in reason as reason’s self-preservation, which is the first methodological principle of reason to orient itself in the supersensible realm. Criticizing Moses Mendelssohn’s position in the controversy in ‘What Does It

Mean to Orient Oneself in Thinking?’, Kant also takes the principle of reason’s self-preservation to bear methodological rather than dogmatic authority. From Kant’s positions in these discussions, I take it that the methodological principle of reason’s self-preservation is best understood as the discipline of pure reason. This principle performs the same function that Kant assigns to the discipline of pure reason in *Critique of Pure Reason*: it preserves reason by preventing its dialectical errors; it is also foundational for reason’s faith (*Vernunftglaube*), which makes the practical use of reason systematic.

I. Kant’s Methodological Reframing of the Question of the Relation between Faith and Reason

The pantheism controversy was arguably the most important controversy among eighteenth-century German intellectuals. In 1835, about half a century after the controversy, Heinrich Heine described its tremendous influence on German intellectual life in this way: ‘Pantheism is the clandestine religion of Germany’ (Heine 2007: 59). Although the controversy started as a dispute between Jacobi and Mendelssohn about Gotthold Lessing’s philosophical views, it became an occasion to raise the old question of the relation between faith and reason in the new historical context of *Aufklärung*—the Enlightenment. In their multiple conceptions of *Aufklärung*, its advocates assumed that reason was the universal authority for all human affairs. The advancement of reason as the universal authority for all human affairs involved far-reaching religious and political implications that seemed threatening to many, especially those holding political power and those who believed in the inviolable authority of religious traditions. In this context, Jacobi’s case against the universal authority of reason captured the attention of opponents of *Aufklärung*. Jacobi highlighted what he took to be an inevitable and troublesome logical conclusion of the universal authority of reason: the denial of freedom and possibility of faith in God. He raised anew the old question of the relation between faith and reason: Given the universal authority that *Aufklärung* grants to reason, is there any room for freedom and faith in God? (For histories of the controversy, see Altmann 1973: 553–759; Beiser 1987: 44–126; Di Giovanni 1992: 3–167)

To address this central question of the pantheism controversy, Kant reframes it *methodologically*, as the question of *how reason should be used* in thinking supersensible objects. The first indication of this methodological reframing is the title of his essay: ‘What Does It Mean to Orient Oneself in Thinking?’ The title implies that the question of the relation between faith and reason essentially belongs to human reason itself. The question concerns how reason should relate to and use itself in thinking supersensible objects. Thinking in the supersensible realm cannot rely on any data or direction from supersensible objects. These objects cannot be given to us and their existence cannot be verified, although this does not mean that they cannot be thought and their existence cannot be presupposed. Reason can think supersensible objects and justifiably presuppose their existence only if it relies on itself to orient itself toward them—that is, only if it proceeds methodologically.

To appreciate Kant's methodological reformulation of the question of the relation between faith and reason, it is useful to distinguish between *the methodological should*, which belongs to Kant's critical methodology, and *the moral should*, which belongs to his moral philosophy. This distinction between systematicity and morality in critical philosophy is neglected in interpretations of 'What Does It Mean to Orient Oneself in Thinking?' (see, for example, O'Neill 1989: 26). *The methodological should* concerns how one should use reason in thinking in order to be systematic whereas *the moral should* is about what one should do in order to be moral. *The methodological should* refers particularly to the principles of reason that make its use (moral or otherwise) systematic or its cognitions complete. Kant takes the methodological principles of reason to belong solely to 'The Doctrine of Method', which is the second part of *Critique of Pure Reason* (1781) in which 'the formal conditions of a complete system of pure reason' are determined (Kant 1997: 627 [KrV A708/B736]). These principles of systematicity are not theoretical; they are not concerned with cognizing what can exist. They are methodological and function to make *all* possible uses or cognitions of reason systematic. From a systematic standpoint, these principles precede the division of reason into two metaphysical (theoretical and practical) uses. Kant notes that the principles outlined in 'The Doctrine of Method' of *Critique of Pure Reason* are foundational for the systematicity of not only the theoretical use but also the practical use of reason: 'The *doctrine of the method* of pure *practical* reason cannot be understood as the way to proceed (in reflection as well as in exposition) with pure practical principles with a view to scientific *cognition* of them, which alone is properly called method elsewhere in the *theoretical* (for popular cognition needs a *manner* but science a *method*, i.e., a procedure in *accordance with principles* of reason by which alone the manifold of a cognition can become a *system*)' (Kant 2015: 121 [KpV 5:151]).

Kant's methodological approach to the question of the relation between faith and reason becomes more explicit in the opening paragraph of 'What Does It Mean to Orient Oneself in Thinking?': 'many *heuristic* methods of thinking perhaps lie hidden in the experiential use of our understanding and reason; if we carefully extract these methods from that experience, they could well enrich philosophy with many useful maxims even in abstract thinking' (Kant 2001: 7 [WDO 8:133]). Kant indicates that by 'abstract thinking' he means the use of reason to think supersensible objects. He reminds the reader that in the pantheism controversy, Mendelssohn attempted to extract the method of using reason 'in abstract thinking': 'Of this kind [of 'maxims. . . in abstract thinking'] is the principle to which Mendelssohn expressly subscribed for the first time. . . in his last writings. . . namely, the maxim that it is necessary to orient oneself in the speculative use of reason' (Kant 2001: 7 [WDO 8:133]). Although Kant does not hold that Mendelssohn correctly understands the method of reason's self-orientation in the supersensible realm, he affirms that the relation between faith and reason should be determined through a principle of reason that orients reason in the supersensible realm.

Thus, the question 'what is the relation between faith and reason?' is reframed as the question 'what is the principle according to which reason orients itself in its

speculative use in the supersensible realm?' This methodological reformulation sets up the framework in which Kant presents the principle of reason's self-preservation. This principle and its methodological nature can be viewed more clearly by examining how Kant criticizes Jacobi's and Mendelssohn's responses to the question of the relation between faith and reason.

2. Kant's Criticism of Jacobi: The Freedom to Think as the Rule of Law in Reason

Kant's criticism of Jacobi focuses on the concept of *the freedom to think* or *intellectual freedom*. Addressing Jacobi and his supporters, Kant writes: 'Without doubt you want to preserve inviolate the *freedom to think* [*Freiheit zu denken*] for without that even your own free flights of genius would soon come to an end. Let us see what would naturally become of this freedom of thought if a procedure such as you are adopting should get the upper hand' (Kant 2001: 16 [WDO 8:144]).

In 'Concerning the Doctrine of Spinoza in Letters to Herr Moses Mendelssohn' (1785), Jacobi argues that reason's demand for complete determination leaves no room for the freedom to think. He writes: 'for the determinist, if he wants to be consistent, must become a fatalist' (Jacobi 1995: 187). Thus, the freedom to think requires refusing the universal authority of reason and its demand for complete determination. For Jacobi, the freedom to think requires freedom from the constraint of the laws of reason. In *Prolegomena to Any Future Metaphysics* (1782), Kant does not deny that reason demands complete determination: 'a critique . . . is never trustworthy unless it is *entirely complete* down to the least elements of pure reason . . . in the sphere of this faculty one must determine and settle either *all* or *nothing*' (Kant 2004: 13 [Prol 4:263]). But Kant disputes that reason's demand for completeness violates the freedom to think. He views Jacobi's position as an attack on reason that undermines the freedom to think. Kant seeks to demonstrate that not only does Jacobi's position fail to preserve the freedom to think, but it is bound to destroy that freedom because preserving the freedom to think requires the constraint of the laws of reason. Kant takes three steps to define what the freedom to think is and how it requires the constraint of the laws of reason. The first two steps identify two types of external constraint in contrast to which Kant defines the freedom to think. The third step focuses on Jacobi's claim to demonstrate that the internal constraint of the laws of reason is fundamental for the freedom to think and affirmatively defines that freedom.

In the first step, Kant defines the freedom to think negatively, by contrasting it with the civil constraint that political authorities might impose: 'The freedom to think is opposed *first* of all to *civil constraint*' (Kant 2001: 16 [WDO 8:144]). Kant explains this point by addressing a conventional objection to it: 'Of course it is said that the freedom to *speak* or to *write* could be taken from us by a superior power, but the freedom to *think* cannot be' (2001: 16 [WDO 8:144]). The reasoning underlying this objection seems to be as follows: communication (speaking or writing) is an external process that external powers can have access to, therefore, the freedom to communicate can be taken from us; on the contrary, thinking is an internal process to which no external power can have access,

therefore, the freedom to think cannot be taken from us. This reasoning questions Kant's opposition between the freedom to think and civil constraint: given that civil constraint is external and thus cannot access our thinking, which is internal, how can it interfere with our thinking and oppose the freedom to think? To indicate how civil constraint is opposed to the freedom to think, Kant responds: 'Yet how much and how correctly we *think* if we did not think as it were in community with others to whom we *communicate* our thoughts, and who communicate theirs with us! Thus one can very well say that this external power which wrenches away people's freedom publicly to *communicate* their thoughts also takes from them the freedom to *think*' (Kant 2001: 16 [WDO 8:144]).

In this passage, Kant does not claim that civil constraint can effectively prevent thinking altogether. Rather, he takes issue with defining the freedom to think too narrowly. Kant implies that the freedom to think is much more than mere isolated thinking of separate individuals, and that it should also take into account 'how much' and 'how correctly' we think. If we do so, we can see how civil constraint violates the freedom to think. In restricting the freedom to communicate our thoughts, civil constraint violates the freedom to think, because it prevents us from being in community and communication with others, and thereby drastically limits how much and how correctly we think. Of course, Kant knows well that an individual might think extensively and correctly without being always physically in community and communication with others. But this does not mean that others are absent in our thinking. Even when we are physically isolated from or non-contemporaneous to others, we think 'as it were in community with others' (Kant 2001: 16 [WDO 8:144]). Thinking may unfold as an internal process in an individual, but thinking is a collective human undertaking. To think, we must use reason, and, in using reason, we ineluctably assume and participate in the intellectual community of all possible users of human reason. This conception of the intellectual community of all possible users of reason is not an external conception of intellectual community, although it inevitably takes an external form in time and space. The intellectual community of all possible users of reason is transcendental-methodological. By virtue of having human reason, one is a possible user of reason, and inescapably participates in the intellectual community of all possible reasoners. This transcendental-methodological conception of intellectual community is usually interpreted in primarily external terms (see, for example, O'Neill 2015: 63; Deligiorgi 2005: 64–65). Hence, if by virtue of using reason we are in an intellectual community with all possible users of reason, the question becomes *how* should we use reason in thinking, or *how* should we relate to others and participate in this intellectual community? Do we use reason, or relate to others and participate in reasoning, in a way that preserves reason as the intellectual community of all possible users of reason or not? This is a *methodological* question that emerges more explicitly after Kant defines the freedom to think negatively and clears the way for his affirmative account in the third step. In the first step of his reasoning against Jacobi, Kant only emphasizes that civil constraint restricts intellectual relation with others in the intellectual community of all possible users of reason to take its external form in time and space. In doing so, it obstructs the development of reason and thus the

improvement of the human condition. This is how civil constraint opposes the freedom to think: it negatively affects how much and how correctly we think, and ultimately how much we improve our ways of thinking (using reason) so as to move toward the ideal intellectual community of all possible users of reason. Kant concludes his first step in the explanation of the freedom to think with a description of the freedom to think as 'that single gem remaining to us in the midst of all the burdens of civil life, through which alone we can devise means of overcoming all the evils of our condition' (Kant 2001: 16 [WDO 8:144]).

In the second step of his account of the freedom to think, Kant presents the freedom to think by contrasting it with a type of constraint that religious authorities might impose on conscience: 'Second, freedom to think is also taken in a sense in which it is opposed to *constraint over conscience*' (Kant 2001: 16 [WDO 8:145]). Here Kant argues against religious restrictions on the freedom to think. As discussed above, civil constraint shapes civil laws to restrict how much and how correctly we think. In distinction, constraint over conscience invokes religious commands to prevent citizens from thinking for themselves. Constraint over conscience can be set up by religious authorities: 'even without having external power, some citizens set themselves up as having the custody of others in religious affairs' (Kant 2001: 16 [WDO 8:145]). Writing in 'An Answer to the Question: What is Enlightenment' (1784), Kant views such guardianship 'in *matters of religion*' as 'the most harmful. . . [and] the most disgraceful of all' (Kant 1996a: 21 [WA 8:41]). The guardians of religious constraint over conscience prohibit reason's free and public examination, 'and instead of reasoning they know how to ban every examination of reason by their early influence on people's minds' (Kant 2001: 21 [WDO 8:145]). They wish to replace citizens' independent thinking with 'prescribed formulas of faith accompanied by the anxious fear of *the dangers of one's own investigation*' (2001: 16 [WDO 8:145]).

Thus, in the first two steps of his account of the freedom to think, Kant defines his conception of the freedom to think negatively, in opposition to civil and religious constraints on thinking. He does so because these constraints restrict the universal authority of reason in civil and religious spheres. Although Kant does not explicitly say so, his reasoning implies that, by opposing *Aufklärung* and its demand for the ultimate authority of reason in *all* spheres, Jacobi and his supporters side with civil and religious constraints and against the freedom to think. The most essential difference between Kant's and Jacobi's conceptions of the freedom to think, however, comes into view most notably in the discussion of another type of constraint that Kant unequivocally reasons for and Jacobi explicitly argues against. As Kant writes two years earlier, we must distinguish between two types of constraint on freedom: 'Everywhere there are restrictions on freedom. But what sort of restriction hinders *Aufklärung* and what sort does not hinder but instead promotes it?' (Kant 1996a: 18 [WA 8:37]). In the first and second steps of his account of the freedom to think, Kant identifies civil and religious constraints as types of restriction that hinder *Aufklärung* and its demand for the universal authority of reason. In the third step, Kant identifies the type of constraint that promotes *Aufklärung* and the universal authority of reason. In contrast to Jacobi's conception of the freedom to think as freedom from the

constraint of the laws of reason, Kant argues that the constraint of the laws of reason is the freedom to think. As will be discussed, it is the freedom to think, or the constraint of the laws of reason, that he ultimately articulates as ‘the maxim of reason’s *self-preservation*’ (Kant 2001: 18 [WDO 8:146]).

In his final step, Kant affirmatively defines his conception of the freedom to think: “**Third**, freedom in thinking signifies the subjection of reason to no laws except *those which it gives itself*; and its opposite is the maxim of a **lawless use** of reason (in order, as genius supposes, to see further than one can under the limitation of laws)’ (Kant 2001: 16 [WDO 8: 145]). Contrary to most interpreters of Kant’s conception of the freedom of think (see, for example, O’Neill 1989; Deligiorgi 2005), this passage does not define the freedom to think primarily as an external (civil or political) form of freedom or right. The freedom to think primarily concerns an internal, particularly methodological, question, that is, the unformulated question in response to which Kant presents his own and Jacobi’s conceptions of the freedom to think: Is reason used lawfully, i.e., according to the maxim of a **lawful use** of reason, or lawlessly, i.e., according to ‘the maxim of a **lawless use** of reason’ (Kant 2001: 16 [WDO 8:145])? This question is the methodological question of how reason should be used in thinking, regardless of what reason is used for. In response to this question, Kant defines his conception of the freedom to think as a way or method of using reason that subjects or constrains reason to its own laws and in so doing enables the lawful use of reason. In short, the freedom to think is the subjection of thinking to laws of reason. Given that reason can *systematically* cognize its own laws and *systematically* subject itself to them only through a critique of reason, we can say that for Kant the freedom to think, or the systematic subjection of reason to its own laws, is the subjection of reason in all its uses to critique. After stating his conception of the freedom to think, Kant presents Jacobi’s opposing response to the same methodological question: ‘if reason will not subject itself to the laws it gives itself, it has to bow under the yoke of laws given by another, for without any law, nothing . . . can play its game for long. Thus the unavoidable consequence of *declared* lawlessness in thinking (of a liberation from the limitations of reason) is that the freedom to think will ultimately be forfeited’ (Kant 2001: 17 [WDO 8:145]). Hence, Kant identifies the freedom to think with *the method* of using reason that subjects reason to its own laws and thus enables reason to be used *systematically* lawfully. The opposing (Jacobi’s) way of using reason is lawless, even when it incorporates some form of law. Kant mentions that lawless uses of reason might incorporate some form of law into themselves in order to attain some stability. A case of the incorporation of some form of law into lawfulness is ‘**superstition**’ which ‘is the complete subjection of reason to facts . . . because this at least has the *form of law* and so allows tranquility to be restored’ (Kant 2001: 17 [WDO 8:145]). *The method* of lawful use of reason is distinguished from the incorporation of some form of law into lawlessness by its *systematic* nature. This method is concerned with the systematic lawfulness of reason. In subjecting reason to its own laws, the method of the lawful use of reason *systematically establishes the rule of law in reason*, prior to the application of any laws of the faculties of reason. Thus, the systematic establishment of the rule of law in reason is the primary condition for the

systematic use of the faculties of reason. It functions as the primary criterion for assessing whether reason is used in a way that can be systematically lawful and thus part of a complete system of reason.

3. The Rule of Law in Reason as Reason's Self-Preservation

Kant's methodological conception of the freedom to think as the method of using reason that subjects reason to its own laws and thus establishes the rule of law in reason indicates that in each and every use of reason the freedom to think or the rule of law in reason is at stake. In each and every use of reason, we respond to the quintessential methodological question: how should reason be used, according to the rule of law or not? To demonstrate this crucial stake in using reason, especially in the supersensible realm, Kant states his conclusion about the consequences of Jacobi's position: 'freedom in thinking finally destroys itself if it tries to proceed in independence of the laws of reason' (Kant 2001: 18 [WDO 8:146]). The stake in Kant's statement will stand out more clearly if we notice that with the destruction of the freedom to think or the rule of law in reason, reason will lack universal authority in all spheres of human existence. If reason cannot establish the rule of law in itself, how can it be the legitimate arbiter of all human affairs? Reason's self-rule is the primary systematic condition of reason's universal authority in all spheres. Hence, at stake in each and every use of reason is the freedom to think, or the rule of law in reason, and at stake in the freedom to think is the very existence of reason as the universal authority in all human affairs. In *Critique of Pure Reason*, in the second section of 'The Discipline of Pure Reason' (Kant 1997: 643 [KrV A738/B766]), Kant emphasizes this point: 'Reason must subject itself to critique in all its undertaking, and cannot restrict the freedom of critique [or reason's self-subjection to its own laws which founds the rule of law in reason] through any prohibition without damaging itself' (Kant 1997: 643 [KrV A738/B766]). Kant makes this essential point more forcefully later in the same paragraph: 'The very existence of reason depends upon this freedom [of critique, or reason's self-subjection to its own laws], which has no dictatorial authority' (Kant 1997: 643 [KrV A738/B766]). As Kant indicates here, reason's self-subjection to its own laws concerns *all* uses of reason, even though in 'What Does It Mean to Orient Oneself in Thinking?', he speaks of this self-subjection as the principle of reason's self-preservation and only with regard to the use of reason in the supersensible realm.

The crucial significance of the freedom to think, or the rule of law in reason, for the very existence of reason as the self-legislative (methodologically autonomous) universal authority in all spheres of human existence leads Kant to use the language of 'reason's *self-preservation*' to present his methodological definition of the freedom to think as the rule of law in reason (Kant 2001: 18 [WDO 8:146]). In each and every use of reason in thinking supersensible objects, reason uses itself in a way, or according to a method, that *either* destroys itself *or* preserves itself as universal authority. Therefore, it is absolutely necessary to determine how reason can preserve itself as universal authority, or how reason can *use* itself without violating its 'holy' and 'original right' as the one and only 'judge' for all human

affairs (Kant 1997: 650 [KrV A752/B780]). Kant addresses this question through a brief explanation of the principle of reason's self-preservation in the last footnote of his essay:

Thinking for oneself means seeking the supreme touchstone of truth in oneself (i.e., in one's own reason); and the maxim of always thinking for oneself is *Aufklärung*. Now there is less to this than people imagine when they place *Aufklärung* in the acquisition of *information*; for it is rather a negative principle in the use of one's faculty of cognition, and often he who is richest in information is the least enlightened in the use he makes of it. To make use of one's own reason means no more than to ask oneself, whenever one is supposed to assume something, whether one could find it feasible to make the ground or the rule on which one assumes it into a universal principle for the use of reason. This test is one that everyone can apply to himself; and with this examination he will see superstition and enthusiasm disappear, even if he falls far short of having the information to refute them on objective grounds. For he is using merely the maxim of reason's *self-preservation*. (Kant 2001: 18 [WDO 8:146])

This footnote suggests that the principle of reason's self-preservation in the supersensible realm consists in reason's subjection to the rule of its own law: to preserve itself in thinking supersensible objects, reason must submit to the rule of law in reason. To explain this account of the principle of reason's self-preservation, the footnote invokes Kant's 1784 definition of *Aufklärung*. Accordingly, *Aufklärung* is 'to make use of one's *own* understanding [reason]' (Kant 1996a: 17 [WA 8:35]), or 'to think for oneself' (1996a: 18 [WA 8:36]). The use of one's own reason is a use in which one's reason orients itself 'without direction from another' (Kant 1996a: 17 [WA 8:35]). This invocation of the definition of *Aufklärung* to explain the crucial significance of the rule of law in reason for its self-preservation in the supersensible realm is directed at Jacobi's and his supporters' hostility to *Aufklärung*, although Kant does not explicitly say so. The invocation indicates that Kant's definition of *Aufklärung* and his definition of the freedom to think are fundamentally the same: the use of one's own reason is essentially one and the same as reason's subjection to its own laws. Only in subjecting one's reason to its own laws, only through the rule of law in reason, can one avoid following the 'direction from another' (Kant 1996a: 17 [WA 8:35]), make reason one's *own*, and use one's own reason. Of course, reason's self-ownership or self-possession should not be understood in a way that substantializes the self or individual human reason. Kant does not hold such a conception of self-ownership. Although in some places Kant associates the language of *possession* with Locke (see, for example, Kant 1997: 221 [KrV A87/B119]), he also uses this language in a methodological sense to articulate his own position. Such use precedes his 1784 essay on enlightenment, where he explicitly speaks of the 'use of one's *own* understanding' (Kant 1996a: 17 [WA 8:35]). In

the preface to the first edition of *Critique of Pure Reason*, Kant characterizes the critical method of reason as 'nothing but the inventory of all we possess [*Besitze*] through pure reason, ordered systematically' (Kant 1997: 104 [KrV A xx]); see also Kant 1997: 645 [KrV A743/B771]). In short, Kant implies that the fundamental identity of reason's individual self-ownership and the universal rule of law in reason exposes the falsity of Jacobi's claim that the universality of reason cancels out the individual character of its use and makes the use of reason necessarily impersonal and contrary to the possibility of individual freedom and personal faith.

The rest of the footnote builds on this basic identity of reason's individual self-ownership and the universal rule of law in reason to sketch whether this identity is realized 'whenever one is supposed to assume something' in the supersensible realm (Kant 2001: 18 [WDO 8:146]). Accordingly, in the speculative use of reason to think supersensible objects, we are bound to answer the methodological question of how to use reason. No matter what the specific assumption is, there are only two possible, indeed opposite, answers: *either* reason is used according to a maxim of lawless use of reason *or* reason is used according to a maxim of lawful use of reason; more specifically, *either* reason is used in a way that 'the ground or the rule on which one assumes it [a supersensible object]' *cannot* be made 'into a universal principle for the use of reason' *or* reason is used in a way that 'the ground or the rule on which one assumes it [a supersensible object]' *can* be made 'into a universal principle for the use of reason' (Kant 2001: 18 [WDO 8:146]). In the first case, one does not 'make use of one's own reason' because one uses reason in a way that does not accord with the rule of law in reason. This way of using reason is not the use of one's own reason as it relies on a ground or rule that one's own reason cannot supply. Using reason to think supersensible objects in this way tends to destroy the universal authority of reason and ultimately leads to enthusiasm. In the second case, one 'make[s] use of one's own reason' because one uses reason in a way that is according to the rule of law in reason. This method of using reason is the use of one's own reason, because through it reason itself can supply the ground of or the rule governing its assumption regarding a supersensible object. This method prevents reason from falling into enthusiasm, preserves the universal authority of reason, and prepares reason 'for *orienting* itself in thinking, solely through reason's own need, in that immeasurable space of the supersensible, which for us is filled with dark night' (Kant 2001: 10 [WDO 8:137]).

Thus, Kant's explanation of the principle of reason's self-preservation as reason's subjection to its own laws implies that the principle of reason's self-preservation is the first methodological principle of reason to orient itself in thinking supersensible objects. Reason can preserve and orient itself in thinking supersensible objects by first establishing the rule of law in reason, that is, by using the method of reason that systematically subjects reason to its own laws. This explanation, however, does not sufficiently elaborate the methodological sense of this principle. In particular, it is not completely clear what the rule of law in reason means for the way the authority of reason in the supersensible realm should be understood. Kant's criticism of Mendelssohn's response to the question of the relation between faith and reason can be helpful in this regard.

4. Kant's Criticism of Mendelssohn: The Methodological Authority of Reason in the Supersensible Realm

In criticizing Jacobi's claim that reason lacks any authority in the supersensible realm, Kant broadly concurs with the position 'which Mendelssohn affirmed, staunchly and with justified zeal' (Kant 2001: 8 [WDO 8:134]). According to this position, 'it was in fact *only* reason—not any alleged sense of truth, not any transcendent intuition under the name of faith' that orients the speculative use of reason in thinking supersensible objects (Kant 2001: 8 [WDO 8:134]).

Although Kant interprets Mendelssohn in a manner and writes about his position in a tone that largely emphasizes their shared views in the controversy, he parts ways with Mendelssohn when it comes to the nature of the authority of reason in its speculative use in the supersensible realm. Kant argues that this authority, which belongs to the principle of reason's self-preservation, is of a methodological not objective nature. On Kant's view, Mendelssohn adopts an 'ambiguous position' (Kant 2001: 7 [WDO 8:134]) on the nature of the authority of reason in the supersensible realm. Kant even goes further to suggest that this ambiguity indicates that 'Mendelssohn himself misunderstood it [the authority of reason in the supersensible realm] for a judgement of *reason's insight*' (Kant 2001: 13 [WDO 8:140]). Kant argues that Mendelssohn privileges an apparently experience-based conception of 'common sense' (Mendelssohn 2012: 73) to orient reason in its speculation beyond experience. Without clarifying the source and meaning of his conception of common sense, Mendelssohn extends its role into the supersensible realm. In Kant's assessment, this extension conflates the sensible realm of possible experience and the supersensible realm beyond experience, which is destined to lead to dialectical errors: Mendelssohn trusted his conception of common sense 'very much in respect of the cognition of supersensible objects, even so far as claiming for it the evidence of demonstration' (Kant 2001: 7 [WDO 8:133]). Similarly, in 'Some Remarks on Ludwig Heinrich Jakob's Examination of Mendelssohn's Morning Hours' (1786), Kant endorses his former student's (Ludwig Jakob's) criticism of Mendelssohn, and questions 'the confidence of this experienced philosopher in the demonstrative method of proof for the most important propositions of pure reason' (Kant 2011: 178 [EBJ 8:151]). The use of the demonstrative method in the supersensible realm is dogmatism. And Kant argues for having 'dogmatism's wings clipped through strict critique' (2011: 178 [EBJ 8:151]).

Kant's case against the adoption of the dogmatic method in the speculative use of reason in the supersensible realm goes back to *Critique of Pure Reason*. In the essay, Kant reminds the reader that '[t]he *Critique* completely clips dogmatism's wings in respect of the cognition of the supersensible objects' (Kant 2001: 15 [WDO 8:144]). In 'The Discipline of Pure Reason,' Kant argues that the speculative use of reason should not rely on the dogmatic method—that is, the method that is based on 'direct synthetic proposition[s] from concepts' (Kant 1997: 642 [KrV A736/B764]). Such propositions presuppose the possible givenness of objects, which cannot be legitimately assumed in the supersensible realm. Kant suggests that the authority of reason in the supersensible should be conceived according to

the systematic method, which relies only on 'reason itself' and does not assume the possible givenness of objects: 'Now if in the content of the speculative use of pure reason there are no dogmata at all, then any **dogmatic** method . . . is inappropriate per se. For it merely masks mistakes and errors, and deceives philosophy, the proper aim of which is to allow all of the steps of reason to be seen in the clearest light. Nevertheless, the method can always be **systematic**. For our reason itself (subjectively) is a system' (Kant 1997: 643 [KrV A737–38/B765–66]). As Kant suggests, the speculative use of reason in the supersensible realm can be legitimate only if it is guided by the systematic method of reason. Kant locates this method in 'reason itself'—that is, reason understood 'subjectively' or in a subjective way or sense. This subjective sense of reason is neither psychological nor objective. Rather, it is *systematic* or *methodological*. In this sense, reason is directed to itself in order to make its concept complete and systematic prior to its uses. This interpretation of Kant's conception of subjective sufficiency as the sufficiency of a systematic or transcendental-methodological principle of reason differs from psychologistic and objective types of interpretation. For example, Lawrence Pasternack and Desmond Hogan develop such interpretations. Pasternack offers a psychologistic interpretation: 'Subjective sufficiency, rather than being about the grounds for assent, is best seen in terms of our psychological orientation toward a proposition. More precisely, when an assent is subjectively sufficient, we may understand this to be a holding-to-be-true with maximal firmness of commitment or confidence, what we sometimes also refer to as 'certainty,' though in its subjective, psychological sense' (Pasternack 2017: 520). Hogan presents some aspects of subjective sufficiency in a way that involves epistemic or objective sufficiency: 'The view that Kant treats the relevant morally-grounded assent as 'non-epistemic' is however mistaken' (Hogan 2009: 60).

Kant elaborates his subjective or transcendental-methodological conception of reason in the second part of *Critique of Pure Reason*, where he lays out the 'formal conditions of a complete system of pure reason' (Kant 1997: 627 [KrV A708/B736])—the methodological conception of reason that also guides Kant's discussion in 'What Does It Mean to Orient Oneself in Thinking?' In this essay, Kant puts his methodological sense of reason similarly to the passage quoted above (Kant 1997: 643 [KrV A737/B765–66]): he speaks of '*only* reason,' 'genuine pure human reason,' 'reason . . . *by itself alone*,' '*reason alone*,' and '*pure* reason' (Kant 2001: 8, 13, 13 [WDO 8:134, 140, 141]). Contrasting 'judgements under a determinate maxim according to objective grounds of cognition' with 'judgements under a determinate maxim according to a subjective ground of differentiation in the determination of its own faculty of judgement,' Kant states that 'to *orient* oneself in thinking in general means: when objective principles of reason are insufficient for holding something true, to determine the matter according to a subjective principle' (Kant 2001: 9–10 [WDO 8:136]). Kant's explanation of these statements implies two closely related but distinct senses of *subjective*. Clarifying these distinct senses of *subjective* can help further illuminate the nature of the authority of reason in the principle of reason's self-preservation. Kant explains his reference to 'a subjective ground of differentiation in the determination of its own faculty of judgement' (Kant 2001: 9–10 [WDO 8:136])

in this way: ‘This subjective means still remaining is nothing other than reason’s feeling of its own need,’ that is, a need ‘attaching to reason in itself’ (Kant 2001: 10 [WDO 8:136]). In a footnote that intends to explain his unusual use of the language of ‘feeling [*Gefühl*]’ for reason, Kant remarks: ‘Reason does not feel; it sees its lack and through the *drive for cognition* [*Erkenntnistrieb*] it effects the feeling of a need’ (Kant 2001: 12 [WDO 8:140]). What Kant generally refers to as ‘reason’s feeling of its own need’ is *subjective* in two closely related but distinct ways: it is subjective in that it is an acknowledgement of reason’s own objective insufficiency. It is also subjective in that it motivates reason to address its objective insufficiency and so methodologically to satisfy its own need in the supersensible realm.

Reason’s acknowledgement of its own objective insufficiency in the supersensible realm is its consciousness of the impossibility of cognizing supersensible objects. In Kant’s account, Mendelssohn’s dogmatic method fails to acknowledge this objective insufficiency. Comparing Mendelssohn’s claims regarding ‘the rational knowledge of God’ (Mendelssohn 2012: 3) to ‘the Cartesian proof of God’s existence,’ Kant writes: ‘*need is taken for insight* . . . so it is also with all the proofs of the worthy Mendelssohn in his *Morning Hours*. They accomplish nothing by way of demonstration . . . Mendelssohn probably did not think about the fact that *arguing dogmatically* with pure reason in the field of the supersensible is the direct path to philosophical enthusiasm’ (Kant 2001: 11 [WDO 8:138]). As Kant points out, without acknowledging reason’s objective insufficiency in the supersensible realm, any presupposition about supersensible objects, including Mendelssohn’s, is bound to degenerate into a dogmatic claim and enthusiasm. After demonstrating this sense of subjective that describes reason’s consciousness of its own objective insufficiency, or its consciousness of the limits of its cognitions, Kant indicates the other way in which reason’s need is subjective. He remarks:

[I]t is not *cognition* but a felt *need* of reason through which Mendelssohn (without knowing it) oriented himself in speculative thinking. And since this guiding thread is not an objective principle of reason . . . but a merely subjective one . . . of the only use of reason allowed by its limits—a corollary of its needs—and since *by itself alone* it constitutes the whole determining ground of our judgement about the existence of the highest being, and its use as a means of orientation in attempts to speculate on this same subject is only contingent, so Mendelssohn erred . . . in that he . . . trusted speculation to the extent of letting it alone settle everything on the path of demonstration. The necessity of the first means could be established only if the insufficiency of the latter is fully admitted. (Kant 2001: 12–13 [WDO 8:139–40])

Here, Kant more clearly distinguishes the two senses or ways in which reason’s need is subjective: (1) reason’s full admission of its own objective insufficiency; (2) reason’s response to its own objective insufficiency—that is, the presupposition of the existence of God. Reason’s admission of its own objective insufficiency indicates the limitation of reason’s theoretical cognition. But this negative orientation of reason in the supersensible realm cannot determine *what* precisely

the need of reason in this realm consists in and *how* it can be satisfied under the rule of law in reason. Thus, reason needs an additional level of orientation. Reason's response to its own objective insufficiency addresses this issue and supplies an additional level of orientation. It determines how, given its objective insufficiency, reason can itself (methodologically) satisfy its 'necessary need' (Kant 2001: 14 [WDO 8:141]) in the supersensible realm. It locates the need of reason in the supersensible realm in its practical use and satisfies it through '*reason's faith*' which 'could be called a *postulate* of reason' (Kant 2001: 14 [WDO 8:141]). The passage also indicates the order of priority between the two ways or senses in which reason's need is subjective. Earlier in the essay Kant states that 'reason's feeling of its own need' can arise 'as a subjective ground for presupposing and assuming something' only if 'reason may not presume to know [supersensible objects] through objective grounds' (Kant 2001: 10 [WDO 8:137]). The passage above reiterates it: Reason's response to its objective insufficiency must presuppose reason's full admission of its own objective insufficiency. In short, the subjective (methodological) principle of reason to presuppose the existence of God can be justifiable only if it presupposes another subjective (methodological) principle of reason: the acknowledgment of the limitation of reason's theoretical cognition.

5. The Disciplinary Nature of the Principle of Reason's Self-Preservation

The distinction between the two senses of subjective, or two methodological principles, in 'What Does It Mean to Orient Oneself in Thinking?' can be traced back to the first two methodological or 'formal conditions of a complete system of pure reason' in 'The Doctrine of Method' in *Critique of Pure Reason* (Kant 1997: 627 [KrV A708/B736]).

The first methodological principle, or the principle of reason's self-preservation, is the discipline of pure reason, because this principle performs the same methodological function that Kant assigns to the discipline of pure reason. In Kant's account, reason's self-discipline consists in reason's acknowledgement of its limitations and consciousness of the necessity to attend and submit to its own laws. This general characterization of the discipline implies both the act of the discipline and the principle that governs this act. Yet Kant does not seem to distinguish between the act and the principle of the discipline. I use the term *discipline* to refer to the principle that governs the discipline rather than the act of the discipline.

The second methodological principle is reason's faith or the postulate of the existence of God, which is implied but not explicitly developed in 'The Canon of Pure Reason'. In *Critique of Pure Reason*, Kant does not use the language of the *postulate* of the existence of God. He uses 'reason's faith [*Vernunftglaube*]' only once (Kant 1997: 689 [KrV A829/B857]). In 'What Does It Mean to Orient Oneself in Thinking?', Kant not only refers to the postulate of the existence of God but also more explicitly indicates what is only implicit in 'The Doctrine of Method' in *Critique of Pure Reason*: the first sense of

subjective, or the discipline of pure reason, is the methodological condition for the second sense of subjective—that is, reason’s faith or the postulate of the existence of God.

5.1 The Principle of Reason’s Self-Preservation Is the Discipline of Pure Reason

The principle of reason’s self-preservation is the discipline of pure reason because it performs the same *negative* methodological function as the discipline: to purify reason and prevent dialectical errors and enthusiasm, through determining the boundaries of the theoretical use of reason.

Kant describes the principle of reason’s self-preservation as ‘a negative principle in the use of one’s faculty of cognition’ (Kant 2001: 18 [WDO 8:146]). Its negativity consists in constraining ‘the *drive for cognition* [*Erkenntnistrieb*]’ that extends itself into the supersensible realm (Kant 2001: 12 [WDO 8:140]) and makes knowledge claims about supersensible objects. This account of the negative function of the principle of reason’s self-preservation mirrors what Kant states in his discussion of negative judgments of discipline in the opening statement of ‘The Discipline of Pure Reason’: ‘In humanity’s general lust for knowledge, negative judgements, which are negative not merely on the basis of logical form but also on the basis of their content, do not stand in high regard: one regards them as jealous enemies of our unremitting drive straining for the expansion of our cognition [*unseres unablässig zur Erweiterung strebenden Erkenntnistriebes*]’ (Kant 1997: 628 [KrV A708/B736]).

In ‘What Does It Mean to Orient Oneself in Thinking?’, Kant also indicates that the negative task of the principle of reason’s self-preservation lies in purifying reason in order to prevent dialectical errors and enthusiasm in thinking supersensible objects. In the supersensible realm, ‘the high claims of reason’s speculative faculty, chiefly its commanding authority (through demonstration) obviously falls away, and what is left to it, insofar as it is speculative, is only the task of purifying the common concept of reason of its contradiction, and defending it against its *own* sophistical attacks on the maxims of healthy reason’ (Kant 2001: 8 [WDO 8:134]). The ‘aim’ of purification ‘is to eliminate the unavoidable dialectic in which pure reason becomes involved and entangled when it is employed dogmatically everywhere’ (Kant 2001: 15–16 [WDO 8:144]). These descriptions of the purifying function of the principle of reason’s self-preservation in ‘What Does It Mean to Orient Oneself in Thinking?’ reiterate central ideas of major passages in the chapter ‘The Discipline of Pure Reason’ (Kant 1997:628–71 [KrV A708–94/B736–822]). Kant writes that in its use in the supersensible realm, reason ‘so badly needs a discipline to constrain its propensity to expansion beyond the narrow boundaries of possible experience and to preserve it from straying and error’ (Kant 1997: 629 [KrV A711/B739]). Such preservation of reason occurs through ‘negative judgements [that] have the special job solely of **preventing error**’ (Kant 1997: 628 [KrV A709/B737]).

In both ‘What Does It Mean to Orient Oneself in Thinking?’ and ‘The Discipline of Pure Reason’ in *Critique of Pure Reason* Kant states that the purification of reason

and prevention of dialectical errors occur through reason's disciplinary determination of its boundaries.

In the essay, Kant says that reason must guard against 'a presumptuous trust in the independence of its faculties from all limitations' (Kant 2001: 17 [WDO 8:146]), which leads to dialectical errors and enthusiasm. And 'there is not a single means more certain to eliminate enthusiasm from the roots up than determination of the bounds of the pure faculty of understanding' (Kant 2001: 15 [WDO 8:144]). The 'single means' is what Kant calls the 'test. . . that everyone can apply to himself, and with this examination he will see superstition and enthusiasm disappear. . . For he is using merely the maxim of reason's *self-preservation*' (Kant 2001: 18 [WDO 8:146]). This negative (preventive) role is also at the center of Kant's discussion of the discipline of pure reason in *Critique of Pure Reason*: the discipline 'does not serve for expansion, as an organon, but rather. . . serves for the determination of boundaries, and instead of discovering truth it has only the silent merit of guarding against errors' (Kant 1997: 672 [KrV A795/B823]).

Although Kant does not use the language of reason's *self-preservation* in *Critique of Pure Reason*, the four sections of 'The Discipline of Pure Reason' (Kant 1997: 628–71 [KrV A708–94/B736–822]) together outline how the discipline preserves pure reason against systemic errors in its use in the supersensible realm. He devotes the first section of the chapter to demonstrating that, unlike the mathematical use of reason, which relies on the construction of concepts within pure intuition, the metaphysical use of reason needs the discipline to delimit the field in which reason can theoretically cognize real objects. Such delimitation enables the dogmatic use of reason in a way that preserves reason against dogmatism. The second section establishes that the disciplinary determination of the boundaries of reason systematically grounds the polemical use of reason, that is, the use of reason in order to defend 'its propositions against dogmatic denial of them' (Kant 1997: 644 [KrV A739/B767]). This delimitation enables the negative (defensive) use of reason in a way that prevents reason from falling into skepticism. The third section presents the discipline as necessary for preventing reason from using a '**transcendental hypothesis**' 'for the explanation of things in nature' (Kant 1997: 660 [KrV A772/B800]). In this way, the discipline preserves reason against 'the appeal to a divine author,' which is 'a principle of lazy reason,' for explaining order and purposiveness in nature (Kant 1997: 660 [KrV A773/B801]). In the fourth (final) section of the chapter, Kant examines three major characteristics of the proofs of the discipline of reason to demonstrate that attending to these characteristics makes reason aware of the impossibility of cognizing supersensible objects and thus protects reason against its dialectical errors.

5.2 The Principle of Reason's Self-Preservation Is the Foundation of Reason's Faith

In addition to negatively orienting reason in the supersensible realm, the negative principle of reason's self-preservation, or the discipline of pure reason, is

foundational for reason's faith. Reason's faith supplies an additional level of orientation for reason in the supersensible realm. It orients reason in its practical use in the supersensible realm to satisfy its need methodologically. In *Critique of Pure Reason*, Kant does not articulate the foundational role of the discipline for reason's faith. But in 'What Does It Mean to Orient Oneself in Thinking?', he discusses this role more explicitly and indicates how the principle of reason's self-preservation, which is the discipline of pure reason, underlies reason's faith. In this way, the discipline is foundational for the systematicity of the practical use of reason. Many readings of Kant's essay do not distinguish these two senses of subjective or the two methodological principles of reason. They do not acknowledge the negative foundation of reason's faith. Such interpretations focus on Kant's account of faith in the chapter 'The Canon of Pure Reason' (Kant 1997: 684–90 [KrV A820–31/B848–59]), without mentioning the foundational role of the reason's self-discipline for reason's faith (for a classical example, see Reinhold 2005: 20; for a contemporary example, see Beiser 1987: 46).

In his discussion of the need of reason in its practical use in the essay, Kant presents a version of his account of the highest good in *Critique of Pure Reason* (Kant 1997: 680 [KrV A810/B839]). The differences in Kant's reasonings for and accounts of the ideal of the highest good in 'The Canon of Pure Reason', 'What Does It Mean to Orient Oneself in Thinking?', and, later, *Critique of Practical Reason*, a discussion of which is beyond the scope of this article, do not affect my claim that the discipline is foundational for reason's faith. The idea that the legitimate postulation of the existence of God, or reason's faith, requires presupposing the discipline of pure reason is fundamental to all variations of Kant's reasoning for the ideal of the highest good. In 'What Does It Mean to Orient Oneself in Thinking?', Kant identifies the need of reason in its practical use and indicates that this need cannot be satisfied by practical reason *itself*. He explains that although 'the practical use of reason consists in the precepts of moral laws', it also *leads* 'to the idea of the *highest good* possible in the world insofar as it is possible only through *freedom*'. But 'these precepts lead to what depends not merely on human freedom but also on *nature*, which is the greatest *happiness*, insofar as it is apportioned according to the first'. To be systematic, in its practical use, 'reason *needs* to assume, for the sake of such a *dependent* highest good, a supreme intelligence as the highest *independent* good. . . only in order to give objective reality to the concept of the highest good' (Kant 2001: 12 [WDO 8:139]). Kant calls this assumption '*reason's faith*, which rests on a need of reason's use with a *practical* intent, [and] could be called a *postulate* of reason' (2001: 14 [WDO 8:141]). Thus, reason's faith orients the practical use of reason in the supersensible realm.

The orientation of the practical use of reason through reason's faith is based on the negative principle of reason's self-preservation. Kant distinguishes '*reason's faith* [*Vernunftglaube*]' from 'a judgement of *reason's insight* [*Vernunftinsicht*]' and 'a judgement from *reason's inspiration* [*Vernunfteingebung*]' (2001: 13 [WDO 8:140]). In Kant's account, both of these judgements presuppose the possible givenness of data from supersensible objects. They do not acknowledge that, in the supersensible realm, reason cannot depend on any data from objects

and must entirely rely on itself to orient itself. In contrast to judgements of 'reason's insight' and 'reason's inspiration' (2001: 13 [WDO 8:140]), 'only reason's faith is one grounded on no data other than those contained in pure reason' (2001: 13 [WDO 8:141]).

Reason's faith requires denying the possibility of any data being given from supersensible objects—that is, denying the possibility of any form of cognition of supersensible objects. He writes: 'All faith is a holding true which is subjectively sufficient, but *consciously* regarded as objectively insufficient' (Kant 2001: 13 [WDO 8:141]). Kant italicizes '*consciously*' to indicate that reason's consciousness of its intrinsic insufficiency to cognize supersensible objects, which in *Critique of Pure Reason* he calls the discipline of pure reason, is foundational for reason's faith. The negative task of the discipline is the condition of the positive doctrine of moral faith.

As Kant writes a year later, it is necessary 'to deny knowledge in order to make room for faith' (1997: 117 [KrV B xxx]). This denial is not based in an extrinsic commitment to Christian faith or liberal toleration. Rather, it is based in an immanent necessity for the formation of reason's faith that enables reason to orient its speculative use in the supersensible realm and thus to satisfy the systematic condition of its practical use. In Kant's view, to 'deny knowledge in order to make room for faith' is to discipline reason, or to acknowledge the objective insufficiency of reason in the supersensible realm, in order to make possible reason's faith as the systematic condition of the practical use of reason. Kant's statement is often interpreted in ways that neglect its primarily methodological meaning. Some do not accept that to 'deny knowledge' in the supersensible realm concerns *all* knowledge of the supersensible (see, for example, Hogan 2009). Others speak of Kant's 'commitment to faith' in an extrinsic or psychological sense (Pasternack 2017: 525). For Kant, to deny knowledge not only preserves reason against dialectical errors and enthusiasm but also functions as the methodological foundation of reason's faith. Only through this denial can faith originate from pure reason and be reason's faith. Kant explains that this denial of knowledge in the supersensible realm, or reason's self-discipline, prevents faith from being transformed into a dogmatic claim: reason's 'faith can never become a knowing through any use of reason' (Kant 2001: 13 [WDO 8:141]). Some interpreters do not acknowledge the methodological nature of the negative judgement that enacts reason's self-discipline and thereby enables reason's faith. They conflate the negative judgement of discipline and reason's faith with judgements of theoretical reason. For example, Allen Wood notes: 'Both faith and knowledge allow us to "assert" (*behaupten*) the judgments which are known or believed true. These judgements presumably are, or may be, theoretical judgements in both cases, and this is why Kant says that "theoretical reason", or the *Erkenntnistriebe*, assumes or presupposes the existence of God and future life on the basis of the moral arguments' (Wood 1970: 15).

Kant's account of the foundational role of the negative principle of reason's self-preservation, or the discipline of pure reason, for reason's faith can be traced back to one of his precritical notes on logic, titled 'on the unchangeability of reason's faith', in *Reflexionen Zur Logik*, note 2446:

It is called moral certainty; but it is not knowledge, but only a faith, as with the jury. If one must absolutely make a judgement due to moral reasons, and must not remain *in suspenso*: then this judgement is necessary.

The principle of self-preservation of reason is the foundation of reason's faith, in which the holding-to-be-true has the same degree of certainty as knowledge, but it is of a different kind, in that it does not rely on the knowledge of reasons in the object, but on the true need of the subject of the theoretical as well as practical use of reason. It always remains to be believed, it will never know, and the former [believing] is the most useful for creatures.—Knowledge puffs up (if it is delusional), but knowledge up to its boundaries (Socrates) makes humble [*Demütig*]. (Kant 1999b [Refl II 16:371–72])

Similar to the last line of this note, the introduction to 'The Discipline of Pure Reason' refers to the discipline as '*humiliation [Demütigung]*' that humbles reason (Kant 1997: 629 [KrV A710/B738]). This humiliation consists in reason's consciousness of its transgression of its own boundaries (Baghai 2019). Kant's note is an early indication of how in 1786 he will articulate his account of the relation between faith and reason in his response to the pantheism controversy: faith is not knowledge; but, to be rational, faith must presuppose—that is, be subjected to—the discipline of pure reason.

Conclusion

The principle of reason's self-preservation in Kant's 'What Does It Mean to Orient Oneself in Thinking?' is best understood as the discipline of pure reason—that is, as the first of the methodological or 'formal conditions of a complete system of pure reason' (Kant 1997: 627 [KrV A708/B736]). This principle negatively orients reason in thinking supersensible objects. It constrains the theoretical use of reason to the sensible realm and thus prevents dialectical errors in the speculative use of reason in the supersensible realm. The principle of reason's self-preservation, or the discipline of pure reason, is also the methodological foundation of reason's faith or the postulate of the existence of God, which in turn systematically conditions the practical use of reason. Given the foundational role of the discipline for systematically grounding both the theoretical and the practical use of reason, Kant maintains that 'the entire philosophy of pure reason is concerned merely with this negative use' (Kant 1997: 629 [KrV A711/B739]), that '[t]he greatest and perhaps only utility of all philosophy of pure reason is thus negative' (Kant 1997: 672 [KrV A795/B823]). Reason can preserve and orient itself in the supersensible realm only if it establishes the rule of law in reason and subjects reason to its own laws. The discipline of pure reason is the foundational institution of the rule of law in reason.

FARSHID BAGHAI 

VILLANOVA UNIVERSITY

farshid.baghai@villanova.edu

References

- Altmann, Alexander. (1973) *Moses Mendelssohn: A Biographical Study*. Tuscaloosa: University of Alabama Press.
- Baghai, Farshid. (2019) 'The Whole of Reason in Kant's Critical Philosophy'. *Dialogue: Canadian Philosophical Review/Revue canadienne de philosophie*, 58, 251–86.
- Beiser, Frederick C. (1987) *The Fate of Reason: German Philosophy from Kant to Fichte*. Cambridge: Harvard University Press.
- Chance, Brian, and Lawrence Pasternack. (2018) 'Rational Faith and the Pantheism Controversy: Kant's 'Orientation' Essay and the Evolution of his Moral Argument'. In Daniel Dahlstrom (ed.), *Aesthetics, History, Politics, and Religion*, vol. 2 of *Kant and His German Contemporaries* (New York: Cambridge University Press), 195–214.
- Deligiorgi, Katerina. (2005) *Kant and the Culture of Enlightenment*. Albany: State University of New York Press.
- Di Giovanni, George. (1992) 'The First Twenty Years of Critique: The Spinoza Connection'. In Paul Guyer (ed.), *The Cambridge Companion to Kant* (New York: Cambridge University Press), 417–448.
- Ferrarin, Alfredo. (2015) *The Powers of Pure Reason: Kant and the Idea of Cosmic Philosophy*. Chicago: University of Chicago Press.
- Heine, Heinrich. (2007) *On the History of Religion and Philosophy in Germany and Other Writings*. Edited by Terry Pinkard. Translated by Howard Pollack-Milgate. New York: Cambridge University Press.
- Hogan, Desmond. (2009) 'How to Know Unknowable Things in Themselves'. *Nous*, 43, 49–63.
- Jacobi, Friedrich Heinrich. (1995) *Concerning the Doctrine of Spinoza in Letters to Herr Moses Mendelssohn (1785)*. In *The Main Philosophical Writings and the Novel 'Allwill'*. Translated by George di Giovanni (Montreal: McGill-Queen's University Press), 173–252.
- Kant, Immanuel. (1901–). *Kants Gesammelte Schriften*. 23 vols. Edited by the Königlich-Preussischen Akademie der Wissenschaften zu Berlin. Berlin: De Gruyter. [I consulted volumes 3, 4, 5, 6, 8, 15, and 16.]
- Kant, Immanuel. (1996a) 'An Answer to the Question: What Is Enlightenment? (1784)'. In Mary J. Gregor (ed. and trans.), *Practical Philosophy* (Cambridge: Cambridge University Press), 17–22.
- Kant, Immanuel. (1996b) *The Metaphysics of Morals*. In Mary J. Gregor (ed. and trans.), *Practical Philosophy* (Cambridge: Cambridge University Press), 363–603.
- Kant, Immanuel. (1997) *Critique of Pure Reason*. Edited and translated by Paul Guyer and Allen Wood. Cambridge: Cambridge University Press.
- Kant, Immanuel. (1999a) *Reflexionen zur Anthropologie. Kants Gesammelte Schriften* 15. Digital edition. Charlottesville: IntelLex Corporation.
- Kant, Immanuel. (1999b) *Reflexionen zur Logik. Kants Gesammelte Schriften* 16. Digital edition. Charlottesville: IntelLex Corporation.
- Kant, Immanuel. (2000) *Critique of the Power of Judgement*. Edited by Paul Guyer. Translated by Paul Guyer and Eric Matthews. Cambridge: Cambridge University Press.
- Kant, Immanuel. (2001) 'What Does It Mean to Orient Oneself in Thinking?'. In Allen W. Wood and George Di Giovanni (eds. and trans.), *Religion and Rational Theology* (Cambridge: Cambridge University Press), 7–18.
- Kant, Immanuel. (2004) *Prolegomena to Any Future Metaphysics That Will Be Able to Come Forward as Science*. Edited and translated by Gary Hatfield. 2nd. ed. Cambridge: Cambridge University Press.
- Kant, Immanuel. (2011) 'Some Remarks on Ludwig Heinrich Jakob's Examination of Mendelssohn's *Morning Hours*.' In *Anthropology, History, and Education*. Edited by Gunter Zoller and Robert B. Loudon. Translated by Mary J. Gregor (Cambridge: Cambridge University Press), 178–81.
- Kant, Immanuel. (2015) *Critique of Practical Reason*. Edited and translated by Mary J. Gregor. Rev. ed. Cambridge: Cambridge University Press.

- Langthaler, Rudolf. (2011) “Das Prinzip der Selbsterhaltung der Vernunft ist das Fundament des Vernunftglaubens”. Eine frühe programmatische These Kants und ihre spätere systematische Entfaltung’. *Internationales Jahrbuch des Deutschen Idealismus*, 7, 58–94.
- Langthaler, Rudolf. (2018) *Kant über den Glauben und die ‘Selbsterhaltung der Vernunft’*. Munich: Verlag Karl Alber.
- Mendelssohn, Moses. (2012) *Last Works*. Translated by Bruce Rosenstock. Chicago: University of Illinois.
- Munzel, G. Felicitas. (1999) *Kant’s Conception of Moral Character: The ‘Critical’ Link of Morality, Anthropology, and Reflective Judgment*. Chicago: University of Chicago Press.
- O’Neill, Onora. (1989) *Constructions of Reason: Explorations of Kant’s Practical Philosophy*. New York: Cambridge University Press.
- O’Neill, Onora. (2015) *Constructing Authorities: Reason, Politics, and Interpretation in Kant’s Philosophy*. New York: Cambridge University Press.
- Pasternack, Lawrence. (2017) ‘Kant on Faith: Religious Assent and the Limits of Knowledge’. In Matthew C. Altmann (ed.), *The Palgrave Kant Handbook* (London: Palgrave-Macmillan), 515–38.
- Reinhold, Karl. (2005) *Letters on the Kantian Philosophy*. Edited by Karl Ameriks. Translated by James Hebbeler. New York: Cambridge University Press.
- Wood, Allen W. (1970) *Kant’s Moral Religion*. Ithaca: Cornell University Press.
- Zammito, John H. (1992) *The Genesis of Kant’s Critique of Judgment*. Chicago: University of Chicago Press.