The Aeneas Argument: Personality and Immortality in Kant's Third Paralogism

Corey W. Dyck

Abstract

In this paper, I challenge the assumption that Kant's Third Paralogism has to do, first and foremost, with the question of personal identity. Beginning with a consideration of the treatments of the soul's personality in Christian Wolff's rational psychology, I show that, despite being influenced by Locke's novel account of personhood and confessing a dissatisfaction with the Scholastic definition of the term, Wolff maintained the agreement between his account of personality and the traditional conception. Moreover, Wolff did not put this concept to a forensic use but considered its primary application to be in the context of the demonstration of the soul's immortality which, according to him, required that after the death of the body the human soul retained its *status personalitatis*, understood as its distinct capacity to be conscious of its identity over time. Wolff's account of the soul's personality, and the use to which he put it, proved rather influential for metaphysicians like G. F. Meier and Moses Mendelssohn, and Kant's lectures in the 1770's also betray this influence. Considering the Third Paralogism in light of this context I claim that, rather than taking up the question of whether the numerical identity of the soul can be inferred from the meagre resources of the *I think*, what is at issue is the rational psychologist's account of how we are conscious of our numerical identity in different times. Despite disagreeing with the rational psychologist on this score, Kant nonetheless contends that the way in which we are, in fact, conscious of our numerical identity in all times qualifies us as persons and suffices for that concept's use in the proof of the soul's immortality. This reading thus makes sense of Kant's claim that the soul's personality, even with its transcendental grounding, is "necessary and sufficient for practical use," and provides a charitable alternative to the recent allegation of a paralogism of pure practical reason on Kant's part.

Horrsecit visu subito, causasque requirit
Inscius Aeneas, quae sint ea flumina porro
Quiue viri tanto complerint agmine ripas?
Tum pater Anchises: animae quibus altera fato
Corpora debentur, Lethaei ad fluminis vndam

Securos latice et longa obliuia potant.
Aeneid, Book VI, as quoted in Meier (1749, 141-2)[1]

Introduction

Kant's Third Paralogism[2] concerns the general topic of the soul's personality (*Personalität*, *Persönlichkeit*), where a *person* is defined in the major of the reconstructed argument as that which "is conscious of the numerical identity of its self in different times" (*CPR* A 361).[3] There is an undeniable resemblance between this definition of person and that offered by Locke which involves, among other elements, a thinking being's capacity to consider "it self as it self, the same thinking thing in different times and places" (1979, 335; II.xxvii.9), and it is perhaps for this reason that the Third Paralogism is frequently understood as dealing specifically with the *identity* of persons. Accordingly, the question at issue is supposed to be the rational psychologist's claim that the soul's consciousness of itself and its identity discloses the numerical identity over time of an immaterial substance. Against the rational psychologist, then, Kant holds that the identity of the *I* of the *I think* is "only a formal condition of my thoughts and their connection, but it does not prove at all the numerical identity of my subject" (*CPR* A 363) and, thus, that in this consciousness of the identical subject of thinking we are not conscious of what is, in fact, our identity as

---

1 In Dryden's translation, the passage reads: "Aeneas wond'ring stood, then ask'd the cause/Which to the stream the crowding people draws./Then thus the sire: 'The souls that throng the flood/Are those to whom, by fate, are other bodies ow'd:/In Lethe's lake they long oblivion taste,/Of future life secure, forgetful of the past'" (1997, 173; lines 964-9).

2 I use 'Third Paralogism' (initial capital letters) when referring to the section as a whole, and 'third paralogism' to refer to the problematic syllogism itself.

3 Translations from the *Kritik der reinen Vernunft* [*CPR*] are taken from Kant (1999), though these have been amended where appropriate. Citations from the *CPR* refer to the pagination in the first "A" edition and, when applicable, to the second, "B" edition. All other citations to Kant's works refer to the volume and page number in the "Akademie Ausgabe" of Kant's *Gesammelte Schriften* [AA] (1900ff). Citations of Kant's works are inserted parenthetically in the text. In addition, I have made use of the following abbreviations in referring to Kant's Nachlass (AA 14-23) and the student notes from his lectures on anthropology (AA 25) and lectures on metaphysics (AA 28-29): *Handschriftlicher Nachlass* [*R*], *Anthropologie-Collins* [*AC*], *Anthropologie-Friedländer* [*AF*], *Anthropologie-Parow* [*APa*], *Anthropologie-Pillau* [*APi*], *Metaphysik L₁* [*ML₁*]. In the case of translations from *ML₁*, I have followed that in Kant (2001); all other translations are my own.

persons. There is, without question, a certain naturalness to this reading as it places Kant's argument at the apex of a line of criticism, beginning with Locke and continuing through Leibniz, of Descartes' inference from the *cogito* to the existence of a thinking substance.[4] Nevertheless, this reading faces a number of textual obstacles,[5] not the least of which is Kant's claim in the concluding paragraph that the concept of personality taken merely in its transcendental signification is "necessary and sufficient for practical use" (*CPR* A 365-6). This would seem to imply that even if the consciousness of our identity expressed in the *I* of the *I think* does not amount to the consciousness of the identity of an immaterial substance, it somehow suffices for what Locke dubbed its "Forensick" use in "appropriating actions and their Merit" (1979, 346; II.xxvii.26). Among the few commentators to confront this confusing result head on, Béatrice Longuenesse has recently contended that an unnoticed "paralogism of pure practical reason" must lie behind Kant's misguided endorsement of this rationalistic conception of person.[6]

Before having recourse to the uncharitable charge of a paralogistic error, however, I think it would be worthwhile to reconsider the assumption that the vast majority of commentators bring to the Third Paralogism, namely, that it has to do, first and foremost, with the identity of persons. Indeed, a glance at the most important treatments of the soul's personality among Kant's predecessors in rational psychology provides a rather different context for Kant's discussion. For them, the personhood of the soul was not understood in terms of its preservation of an appropriate identity over time as disclosed by its consciousness of itself; instead, the human soul was accorded the status of personhood insofar as the very capacity for a consciousness of its own identity served to distinguish it from animal souls. Unsurprisingly, this concept of person was not put to a forensic use, since it had nothing to do with the individuation of human souls, but instead found its exclusive employment in the demonstration of the soul's immortality which required, among other things, that the human soul retain its distinctive *status personalitatis* after the death of the body. It would appear that the question of the soul's personality took on increasing importance through the mid-1700's, and that even Kant himself under-

---

4 For this view, see Allison (1969, 121-2).

5 For instance, as Ameriks (2000, 132) notes, Kant appears to saddle the rational psychologist with an empirical claim (that the soul is consciousness of itself as in time). Similarly, Longuenesse (2007, 157) notes that the footnote in the section has been a perennial source of trouble for commentators.

6 Longuenesse (2007, 161-2).

stood the concept of person in this way in the student notes from his lectures on anthropology and metaphysics from the 1770's. Reading Kant's Third Paralogism in light of this context, it becomes clear that what is primarily at issue is not the putative inference from the *I think* to the soul's identity over time, but rather the *way* in which the rational psychologist alleges that the soul is conscious of its numerical identity and, thus, a person. In spite of disagreeing with the rational psychologist on this score, Kant will claim that the soul's consciousness of its identity expressed in the *I* of the *I think* nonetheless qualifies it as a person and even suffices for that concept's use in the demonstration of the soul's immortality.

1. The Definition and Use of the Concept of Person in Wolff's Rational Psychology

Discussions of the background to the Third Paralogisms often, and rightly, begin with a nod towards Locke's famous (if not unproblematic) definition of the term in the *Essay* as "a thinking intelligent Being, that has reason and reflection, and can consider it self as it self, the same thinking thing in different times and places." I take it, however, that Locke presents an appropriate starting point not because he would exercise a direct influence on Kant's thought in the Third Paralogism, but rather because Christian Wolff would offer a definition of personhood that closely approximates Locke's novel account, though Wolff would adapt it to his own ends in rational psychology.[7] There is an unmistakable resemblance between Locke's treatment and Wolff's account of personhood near the conclusion of the rational psychology of the *Deutsche Metaphysik* (*German Metaphysics*, first edition 1719).[8] There, Wolff claims that "a thing is

7 I am not aware of any evidence of Locke's direct influence on Wolff on the question of personhood. For an appraisal of Locke's general influence on Wolff, see Beck (1969, 267).

8 For ease of reference, citations to Wolff's works are inserted parenthetically in the text, referring to the relevant section number, and I make use of the following abbreviations: *Vernünfftige Gedancken von Gott, der Welt und der Seele des Menschen* (*Deutsche Metaphysik*) [*DM*]; *Vernünftige Gedancken von Gott, der Welt, und der Seele des Menschen, Anderer Theil, bestehend in ausführlichen Anmerckungen* (*Anmerkungen zur Deutschen Metaphysik*)[*AzDM*]; *Ausführliche Nachricht von seinen eigenen Schrifften, die er in deutscher Sprache heraus gegeben* [*AN*]; *Philosophia rationalis, sive Logica* [*LL*]; *Psychologia empirica* [*PE*]; *Psychologia rationalis*

called a *person* that is conscious that it is the very same thing that was previously in this or that state" (*DM* § 924), which definition clearly evokes Locke's original in making the (actual) consciousness of identity essential to personhood. And while Wolff appears to differ from Locke in not additionally taking a person as "a thinking intelligent being," that is, a being with "reason and reflection," he will in fact claim that the consciousness of our own identity at different times already implies the possession of these faculties. The consciousness of our identity at different times, as the distinct representation of the connection of two states, involves the use of reason inasmuch as reason is simply the "insight into the connection of truths," and in this case an insight into how "one [state] is grounded in the other" (*DM* § 865).[9] This use of reason itself requires the operation of the faculty of reflection, defined as the faculty through which we direct our thought to one part of the manifold of a representation after another (cf. *DM* § 272).

It might thus appear that the account of personhood Wolff offers in these passages of the *Deutsche Metaphysik* is reconcilable with Locke's own treatment; nonetheless, the differences between the two become clearer later in the *Psychologia rationalis*, in the course of Wolff's comparison of his conception of person with that of the Scholastics. There, Wolff offers the slightly modified definition of a person as that "being which preserves a memory of itself, that is, which remembers that it is that same being that was previously in this or that state" (*PR* § 741). Given, however, that memory involves the recognition, or consciousness that we have previously had an idea (cf. *PE* §§ 173, 175), this definition does not differ significantly from the previous one.[10] In spite of this Lockean definition, however, Wolff will stress the "agreement [*concordia*]" between his account and the Scholastic definition of person as a *suppositum intelligens* against those who would charge that his conception would establish two distinct persons (a divine and human) in Christ. Taking a *suppositum intelligens* to be a singular, living substance endowed with an intellect, Wolff argues that a being with a memory and which remembers itself to be the same

---

[*PR*]; *Prima philosophia sive Ontologia* [*Ont*], *Jus naturae* [*JN*]. Translations from Wolff's German and Latin texts are my own.

9 That this representation is "distinct" further implies the activity of the understanding; cf. *DM* § 277.

10 This difference is commensurate with the claim of certain commentators that the *DM* is a "modern" philosophical text, setting out as it does from our consciousness of ourselves (and other things), whereas Wolff's Latin works abandon this starting-point and retreat into a kind of Aristotelianism. See Wundt (1945, 152).

being as was previously in some state will necessarily be a singular being because it recognizes itself as the *same* being as before (*PR* § 741); it will also be endowed with an intellect since this recollection involves a distinct cognition, namely the *judgment* that it is the same being (*PR* § 281). Such a being will also constitute a living substance insofar as any being endowed with an intellect will possess a power of cognition (*vi cognoscitiva*; cf. *PE* § 275), where a power can only be found in a substance, and a substance is living in virtue of its possession of such an intrinsic power (cf. *PR* § 741, *Ont* §§ 725, 768). Wolff concludes that a person is in fact "a singular, living substance and so a *suppositum*" endowed with an intellect (*PR* § 741) and, on account of the agreement between these two accounts of personhood, that neither definition is more (nor less) guilty of making Christ into two distinct persons than the other.[11]

Even granting this "agreement," however, we should be wary of taking Wolff to issue a blanket endorsement of the Scholastic conception of person, as Wolff himself views his definition as a distinct improvement over that of his predecessors. In the case of the Scholastic definition, Wolff claims that the *defeniens*, *suppositum intelligens*, is obscure inasmuch as a concept drawn from it will not be sufficient for the recognition of instances that fall under the concept and for distinguishing them from other things (cf. *LL* § 80). In this case, Wolff apparently construes the term *intelligens* as vague because it does not specify appropriate criteria through which the distinction between persons and animals can be recognized.[12] This is not to say that a person is not, ultimately, a *suppositum intelligens*, but only that we do not gain a clear sense of what constitutes a person by means of this definition. Indeed, Wolff will claim that even when we do, in the course of daily affairs, recognize individuals as persons, it is a function of our ordinary use of language and not the clarity of the terms in this definition. As Wolff writes in the *Ontologia*:

> The Scholastics define *existence*, *substance*, *person* badly. These terms are not for this reason obscure, but instead are clear, since to them [i. e., the terms] correspond clear concepts acquired through use [i. e., rather than through the defini-

11 See *AzDM* § 339: "Weil ich sage, eine *Person* sey ein Ding, das sich bewust ist, es sey eben dasjenige, was vorher in diesem oder jenem Zustande gewesen; so haben einige vermeynet, als wenn ich in Christo zwey Personen statuiren müste. Mich wundert dergleichen Schluss: denn man nehme die gemeine Erklärung der Person an, die in allen alten *Compendiis Theologiae* stehet, und in den alten Metaphysicken überall zu finden, dass sie sey *Suppositum intelligens*, so wird man seine *Consequentiam* eben so wohl als bey meiner anbringen können."

12 Compare Wolff's diagnosis of a parallel fault with the Scholastic definition of *mode* in *Ont* § 14.

> tion] in virtue of which we recognize existence, substance, and person. (*Ont* § 13).

What is needed, then, is a definition that enumerates marks that distinguish more clearly between persons and animals, and Wolff naturally takes his own definition to constitute just such an improvement. He notes that the souls of animals lack reason and understanding, which implies that they cannot be aware of their previous states (*DM* § 869) and renders them incapable of any consciousness of the fact that they are the same thing that was previously in this or that state (*DM* § 923; *PR* § 766); thus, the souls of animals cannot be considered persons (*DM* § 924; PR § 767). The souls of human beings are, however, persons since they possess the relevant faculties and are conscious of their identity in this way (*DM* § 892, § 924; *PR* § 767). [13] Wolff's account of the soul's personality is, for the preceding reasons, fairly distinctive. Unlike Locke, he explicitly takes a person to amount to a "singular, living *substance*," but like Locke, he privileges the consciousness of ourselves and our identity as the true mark of personhood.

The significance of Wolff's conception of personality becomes clear in his discussion of the soul's immortality, in the course of which the issue of the soul's personality is originally introduced. Wolff takes the immortality of the soul to involve its survival of the death of the body and, like Descartes,[14] Wolff thinks that this survival involves the soul's incorruptibility, or the fact that the soul does not cease to be through the dissolution of its parts (cf. *PR* § 666). As a simple substance (*DM* § 742; *PR* § 48), the soul lacks parts and thus it could only be destroyed through instant annihilation (cf. *DM* § 922; *PR* § 732) rather than through the gradual dispersal of its parts, and for this reason it must be incorruptible (*DM* § 922; *PR* § 729). Inasmuch as the corruptibility of the body is evident in its putrefaction after death, it follows that the soul cannot perish along with the body (*PR* § 730), and, while a sufficiently powerful being, such as God, could annihilate the soul after the death of the body, there is no reason why this should occur (*PR* § 744). Thus, Wolff concludes that the soul survives,

---

13 Similar reasoning lies behind Wolff's claim that the person is an *individuum morale* (cf. *PR* § 740). Wolff understands by the moral state a condition of being subject to laws and obligations (*JN* I, § 123), for which the possession of reason is required (*JN* I, § 70) and which, therefore, distinguishes the individual human, as a moral being, from animals.

14 See Descartes (1999, 9-10). For his account of immortality, Wolff is likely influenced by Leibniz's treatment in the *Theodicy* (1990, 171), and he also acknowledges a debt to a dissertation of his student, Ludwig Philipp Thümmig (1721); see *DM* § 925, *AzDM* § 341, and Thümmig (1725, 199-200).

and persists long after, the death of the body. Yet, while Wolff takes the soul's survival of the death of the body to be a necessary condition of immortality, he does not think that it is alone sufficient. As he observes, it could be the case that the soul survives the death of the body but is no longer capable of distinct perceptions, a condition likened to the soul's falling asleep (*DM* § 925, *AzDM* § 340) or *psychopannychia* (*PR* § 739), a possibility that contradicts the portrayal of the afterlife in Scripture.[15] Indeed, since Wolff has claimed that the soul is capable of representing the world in accordance with the position of the body with which it is united (cf. *DM* § 753; *PR* § 62), it would seem that once the sensory organs cease to function the soul must lose this capacity, leaving it with only confused and obscure perceptions (*AzDM* § 338). While Wolff thinks that any account of the union of the soul and body in terms of physical influx cannot avoid this result, he argues that the remaining systems in which "the soul has absolutely no need of the body" as a source of its perceptions fare much better (*AzDM* § 340; *AN* § 106). In accordance with his hypothetical endorsement of the system of pre-established harmony,[16] Wolff claims that the soul must be taken to exist previous to its union with the body (cf. *DM* § 789; *PR* § 727), in which condition the soul's sensations have little clarity (*DM* § 925). When the soul comes to be united to the body, even though this constitutes a "great change [*grosse Veränderung*; *magna mutatio*]" (*DM* § 925; *PR* § 745), the soul does not lose its obscure perceptions but rather many of these become clear. From this, Wolff infers that "in great changes the soul retains that which it has, and receives still more than it previously had" (*DM* § 925), and given that the death of the body constitutes yet another great change, Wolff concludes that, far from losing our capacity for distinct perception, there is little reason to doubt that our remaining obscure and confused perceptions will become more clear and distinct upon the body's perishing (*DM* § 925; *PR* § 745).[17]

Even with the soul's survival of the death of the body in a state of distinct perceptions, however, Wolff still thinks that one thing is lacking for genuine immortality, namely, the personality of the soul. It could be the case, for instance, that the soul survives the death of the body and retains the capacity for distinct perceptions but that it is no longer aware of being the same soul now as it had been when, united with the body, it was in a

15 For specific passages, see Calvin (1932) *passim*.

16 For more on this point, and its relevance for Wolff's conception of rational psychology, see Dyck (2009, 254-5).

17 In *AzDM* § 340 and *PR* § 745, Wolff defends himself from the charge that the damned would be unjustly rewarded through this perfection of cognition.

certain state. Here again, Wolff considers the possibility of such a condition to be contrary to the pronouncements of Scripture (*PR* § 740), but he notes as well that it would undermine the justification for punishment (and reward) in the afterlife:

> Since if souls no longer knew that they were the same beings which had committed this or that offense in life, then one would no longer find a reason why they should be punished after death as they would no longer know why this would befall them. (*AzDM* § 341; cf. also *PR* § 740).

Against this, Wolff argues that, given that the soul continues to have distinct perceptions after the death of the body (*PR* § 745), it must also be the case that the soul is conscious of itself after the death of the body, as the states of distinct perception and such consciousness are mutually implicative (cf. *PR* § 13). In addition, given what Wolff calls the law of imagination, or the law that past perceptions are reproduced in virtue of their similarity to present perceptions (*PE* § 117), when the perceptions that we have after the death of the body have something in common with those of our embodied life, the past perception will be reproduced and, by means of our memory, we will recall that we have previously had that perception before (cf. *PE* §§ 173, 175). Thus, after the death of the body, the soul will be conscious that it remains the same soul as it had been previously when, in union with the body, it was in this or that condition, and thus it will maintain its state of personality (*PR* § 746). Insofar, then, as the soul survives the death of the body, remaining capable of distinct perception and its consciousness of itself, it must be immortal (*PR* §§ 739, 747).

Whatever the broad continuities that obtain between Locke's and Wolff's definitions of personality, then, Wolff's employment of the concept differs drastically from that of Locke. Locke had claimed that *person* "is a Forensick Term appropriating actions and their Merit" (1979, 346; II.xxvii.26), and for him the prospect that we might not retain our identity as persons after death posed a threat to divine justice since, in the absence of any continuity of consciousness, an individual would not be the same person as the one who had committed the original action and so would not be justly punished for it (cf. 1979, 347; II.xxvii.26). It might be supposed that Wolff also thought it important to demonstrate that personality obtains after death to avoid implicating divine justice, but this would be mistaken. For Wolff, the concept of person does not bear the weight of moral responsibility since the being to which an act is ultimately attributed is the individual substance of which the state of personality is predicated; consequently, what is at issue in personality is only whether a

given soul *knows* or *recognizes* that it is the same as it was previously, where this knowledge is not somehow constitutive of our identity but presupposes it. As it relates to immortality, then, the important question for Wolff cannot be whether the *same person* remains after death (or in life, for that matter), but rather whether we continue to possess the *status* of persons upon the death of the body, that is, whether we continue in the same state of having a memory of ourselves (and distinct perceptions) as we had been in previously (*PR* § 742).[18] So, Wolff writes that in order to demonstrate immortality "we must further prove that the souls of men also retain their *statum personaliter*, that is, that they remain conscious *that* they are this and that soul that found itself in this or that condition in the unification with the body" (*AzDM* § 338 – my latter emphasis). And while we have seen that Wolff thinks that punishment and reward in the afterlife would be without justification if the soul did not maintain its personality, this claim is founded upon Wolff's conception of the relation of punishment and reward to God's plan in creating the world. According to Wolff, God created the world so that human beings could recognize His perfection, and God's intention in punishing the evil-doer is to reveal His wisdom, which is accomplished through bringing about "an arrangement of things so that from it His hate for the evil conduct of man becomes clear and therefore, notwithstanding His goodness, He lets them come to harm, that is, He punishes evil" (*DM* § 1084). This recognition is, of course, impossible when those who should face punishment, or enjoy reward, in the afterlife do so without a memory of themselves and their actions in their previous condition. Thus, it is God's wisdom, rather than justice, that would be directly implicated by the loss of the soul's state of personality after death, since God would have arranged the world in such a way that His perfection would not be evident to all.

2. Doubts about the Soul's Personality after Death

As might be expected, Wolff's definition of the concept of the soul's personality, and the use to which he put it, proved rather influential, particularly as it relates to the immortality of the soul. This is indicated in the numerous and detailed treatments of the topic that appeared after Wolff's original discussion in the *Deutsche Metaphysik*. Among the first of these is that of Johann Gustav Reinbeck, whose *Philosophische Gedancken über die*

18 For Wolff's definition of a state, see *DM* § 121 and *Ont* § 705.

vernünfftige Seele und derselben Unsterblichkeit (*Philosophical Thoughts on the Rational Soul and its Immortality*, 1739) provides a number of arguments for the soul's immortality, by which he understands the life of the rational soul after death,[19] while remaining faithful to Wolff's general presentation. He argues, for example, that the immortality of the soul can be proved through a demonstration of the soul's personality since, if it can be shown that the soul preserves its memory of itself after death, then it must be conscious of itself as well and so it must continue its rational life.[20] Israel Gottlieb Canz sought, in his *Uberzeugender Beweis aus der Vernunft von der Unsterblichkeit sowohl der Menschen Seele insgemein, als besonders der Kinder-Seelen* (*A Convincing Proof from Reason of the Immortality of the Human Soul in general and of the Souls of Children in particular*, 1744), to improve on Reinbeck's treatment by, among other things, "making evident, with sound reasons, all the partial concepts [*Theilbegriffe*] which are contained in the soul's immortality."[21] Following Wolff, Canz enumerates incorruptibility, the condition of distinct perceptions, and personality[22] as the appropriate partial concepts and, in the section devoted to the soul's personality, sets out to prove that "in accordance with God's decree, the soul will recognize that it is the very same which had acted in such a way in life."[23] Canz is also clear as to why the failure to preserve the state of personality of souls after the death of the body would be inconsistent with God's purpose in creating the world, namely, that God's wisdom should be honoured:

> For that soul which did not hold itself for the very same after death as it was in this life, the goodness and power of God would no doubt be obvious [...] but God's wisdom would not be evident to it since it would not regard its future life as the aim of the current one, and would not know that it had been in this life and through this had been lead to eternal life. The wisdom of the highest is usually expressed with respect to these things.[24]

Another notable treatment is, of course, that found in the third conversation of Mendelssohn's *Phaedon*. Cebes asks Socrates whether the soul "is destined for true immortality, for the eternal endurance of its consciousness and distinct feeling of self, or if these favours of the creator cease

19 Reinbeck (1739, 28).

20 Cf. Reinbeck (1739, 223-231).

21 Canz (1744, 5).

22 Canz (1744, 13-16).

23 Canz (1744, 299).

24 Canz (1744, 306-7). It should be noted that Canz also thinks the soul's personality after death follows from God's goodness (308-9) and justice (310-11).

again after a brief enjoyment and an eternal forgetfulness takes their place."[25] In favour of the preservation of our feeling of self after death, Mendelssohn's Socrates argues that its loss would be inconsistent with divine wisdom since God has made it the end of our natures to strive endlessly to perfect ourselves (a task whose completion requires eternal life), and the feeling of self is essential for this continual perfection of our capacities.[26]

Among the most original treatments, however, is that offered by Georg Friedrich Meier in his *Gedancken von dem Zustande der Seele nach dem Tode* (*Thoughts on the State of the Soul after Death*, 1746). There, Meier contrasts two senses of the soul's immortality: a narrower sense which involves only the soul's continued duration and life after the death of the body, and a broader sense introduced by the moderns which also includes the soul's consciousness of itself and of other things after death, as well as its retention of its states of distinct perception and personality.[27] Meier claims, however, that in whatever sense it is understood, the soul's immortality cannot be known with certainty. For instance, assuming that the soul is a simple, immaterial substance, it follows that the soul does not naturally cease to be,[28] but we cannot be certain that the annihilation of a soul after death does not belong to the best of all possible worlds.[29] As regards the soul's personality, Meier notes that philosophers since Wolff have been uniformly hostile to the conjecture of the "heathen philosophers" that in the next world "there is a river of forgetfulness, from which the departed souls must drink, through which draught the representations of their previous states are utterly washed away."[30] Indeed, Meier agrees with those modern philosophers to the extent that "it is much more natural that the soul should recall its previous life after death,"

> [s]ince as soon as we receive a representation that is similar to a previous one, the latter is brought to light through the former. Now [assuming that God has decided that the soul should have a functioning memory after death], it must be accepted without dispute that after death the soul will receive representations that are similar to those we presently have, because otherwise there would be a leap in

25 Mendelssohn (1979, 105).

26 See Mendelssohn (1979, 106): "Mit der gesamten Reihe der denkenden Wesen hat es die nehmliche Beschaffenheit: so lange sie mit Selbstgefühl empfinden, denken, wollen, begehren, verabscheuen, so bilden sie die ihnen anerschaffenen Fähigkeiten immer mehr aus."

27 Meier (1746, 48-9). Cf. also Meier (1746, 36).

28 Meier (1746, 56-7).

29 Meier (1746, 69-72).

30 Meier (1746, 141).

> the development of its concepts; thus, if the soul should forget itself, then it would have to give up its entire nature as God would have to revoke the law of the imagination.[31]

Yet, because this argument rests on a decision taken by God, into which we can gain insight neither by reason nor experience, Meier concludes that it amounts to a merely *probable* opinion: "It remains uncertain therefore whether the soul retains its personality after the death of the body."[32] And even though Meier acknowledges the importance of this opinion for the purposes of reward and punishment in the afterlife,[33] he thinks that it can be held only with the "certainty of faith" rather than that of reason.[34]

Whether or not Kant was familiar with Meier's treatment, he expresses similar doubts regarding the cognition of the soul's immortality, and its personality after death in particular, in the lectures of the 1770's. However, it is not clear after an initial glance at the student lecture notes from Kant's courses on anthropology and metaphysics that he shares the Wolffian conception of person. For instance, rather than explicitly taking personality to consist in our consciousness of our previous condition, Kant usually ties personality to the ability to think or have the concept of the *I*. Thus, the ML_1 notes read:

> The consciousness of one's self, the concept of the I, does not occur with such beings that have no inner sense; accordingly, no non-rational animal can think: I am; from this follows the difference that beings that have such a concept of the I possess *personality*. (ML_1, AA 28:277).

In addition, Kant distinguishes between two different senses in which this *I* can be understood, that is, the *I* considered *in sensu latiori* and *in sensu stricto*. In the ML_1 notes, the distinction is characterized as follows: "This *I* can be taken in a twofold sense: *I as human being*, and *I as intelligence*. I, *as a human being*, am an object of *inner* and *outer* sense. I *as intelligence* am an object *of inner sense only*" (ML_1, AA 28:224)."[35] The *I* in the broad sense is identified with that representation of the *I* that we have of ourselves insofar as we are also subject to alterations of the body occasioned by external objects, that is, it is a representation of the soul as it is "deter-

31 Meier (1746, 143).

32 Meier (1746, 143-4). See also Meier (1746, 144).

33 Meier (1746, 49-50).

34 On this distinction, see Meier (1746, 8).

35 See also *AC*, AA 25:13, *APa*, AA 25:245-6, and *AF*, AA 25:473-4.

mined by the body, and stand[s] with it in interaction" (*ML*$_1$, AA 28:225).[36] By contrast, the *I* in the narrow sense expresses the consciousness of that self or subject to which all of our representations are related: "Each human as a person or intelligence relates all thought to itself by means of the I" (*AC*, AA 25:476).[37] This twofold significance of the concept of the *I*, culminates in Kant's unusual, and rather un-Wolffian, claim that we have a *double* personality: "Every human being has within him, as it were, a double personality, the I as soul and the I as human being" (*AC*, AA 25:13). This might be taken to suggest that Kant has abandoned (if he had ever adopted) the Wolffian conception of personhood as designating a complex capacity of the soul,[38] and indeed, that Kant should also claim that "personality makes it that something can be imputed [*imputirt*] to me" (*AC*, AA 25:11) would seem to leave little doubt about this.

In spite of these obstacles, the continuity between Kant's and Wolff's accounts can be demonstrated. First, it should be noted that, in linking personality to the concept of the *I*, Kant is actually taking up a remark by Wolff in which it is claimed that a human is recognized as a person through the use of the word *ego*.[39] Second, the doctrine of double personality is not necessarily inconsistent with the Wolffian conception, as long as it is not taken to mean that we are, in fact, two distinct persons but rather only that there are putatively two senses in which we can be considered to be persons corresponding to these two ways in which we might

---

36 On the psychological consideration of the I in the broad sense, see *ML*$_1$, AA 28:259: "When we consider the soul of a human being, we regard it not merely as intelligence, but rather when it *stands in connection with the body* as soul of a human being." See also *R* 5461: "Ich als das *correlatum* aller äusseren Anschauung bin Mensch. Bei äussere Anschauung, worauf sich all übrige an mir beziehen, ist mein Korper. Also muss ich als ein subject äussere Anschauungen einen Korper haben" (AA 18:189).

37 See also *APi*, AA 25:736: "Das Vermögen eines Geschöpfs sich selbst anzuschauen, und alles in der Schöpfung auf sich zu referieren ist die Persönlichkeit."

38 For this objection, consult Ameriks (2000, 171n8). See also *APa*, AA 25:246 and *AF*, AA 25:475-6. The version of this claim in the anthropology lectures perhaps owes something to Kant's previous contrast, evidently inspired by Swedenborg, between the personality of the human being and that of the spirit, as he claims that the subject of such a "dual personality" is not "one and the same person" (cf. AA 2:337-8 and note). In the *APi*, the notes appear to go even further, as they read that "[o]f myself, I find a twofold subject" (AA 25:736).

39 See *PR* § 743: "Nimirum homo qua persona intelligit, quid sit illud quod vocabulo *ego* indigitatur."

become conscious of ourselves.[40] In any case, perhaps seeking to head off this likely confusion, Kant claims that our personality is only properly expressed by the *I* in the narrow sense and he does so for the now-familiar Wolffian reasons. Kant is clear that the *I* in the broad sense is the first thought that occurs to us when we turn our attention to ourselves,[41] whereas the *I* in the strict sense is obtained only by means of abstracting from all that belongs to outer sense in the original representation. Thus, at AA 28:265, the *ML1* notes read that "I take the self in the strict sense insofar as I omit everything that belongs to my self in the broader sense" and, a few lines later, it is claimed that this concept is attained when we "abstract everything outer from the object of inner sense." Significantly, it is this narrower sense of the concept of the *I* that is at issue in rational psychology[42] and, given that it is the concept of that subject that "is no predicate of another thing," the "I, or the soul through which the I is expressed, is a substance" (AA 28:266). On the basis of this substantiality of the *I* in the strict sense, or soul, which contrasts with the mutability of the body, the consciousness of which is expressed by the *I* in the broad sense, Kant contends that the *I* in the strict sense is the "proper" *I*:

> The proper [*eigentliche*] I is something substantial, simple and persistent; on the contrary, one regards the I as human being as changeable – for example, one says *I was large*, *I was small. The I* [i. e., in the narrow sense] *would not change if one was in another body*. (*AC*, AA 25:13).

Thus, Kant, like Wolff, holds that only a substance (indeed, only a simple, immaterial substance) is the proper subject of personality since only then is the consciousness of our subject the consciousness of a being that does not change, or that is strictly speaking the same over time.[43] Finally, when

---

40 Compare Klemme: "Mit dem Ausdruck 'doppelte' Persönlichkeit ist nicht gemeint, dass er wirklich zwei verschiedene Persönlichkeiten hat, sondern dass die eine Persönlichkeit in zwei verschiedenen Hinsichten von ihm ausgesagt werden kann, nämlich als Seele und als Mensch" (96). Kant's introduction of the notion of a *combined* personality might be intended as a clarification along just these lines: "Die Intelligenz macht die Persöhnlichkeit aus, die aber mit der Thierheit combinirt ist" (*AF*, AA 25:476).

41 Cf. *AC*, AA 25:10 and *APa*, AA 25:244.

42 "When I speak of the soul, then I speak of the I *in sensu stricto*. We receive the concept of the soul only through *the I*, thus through the inner intuition of inner sense" (*ML1*, AA 28:265).

43 On this point, see *R* 4237: "Der Mensch hat zweyerley Leben: 1. das thierische, 2. das geistige. Das letztere ist die personlichkeit" (AA 17:471-2), *R* 4240: "Das *principium* des Lebens im Menschen, das eigentliche subiect seiner persöhnlichkeit, sofern es die Materie belebt, heisst seele, und das Leben des Menschen ist ein thierisch Leben" (AA 17:474), and *R* 5049 (AA 18:72).

it comes to Kant's explicit connection between personality and imputation, it must be borne in mind that Kant frequently distinguishes between psychological and moral, or practical, personality (where this is unrelated to the doctrine of double personality). We find the following account of this distinction in the *ML1* notes:

> This [consciousness of one's self, the concept of the *I*] is psychological personality, to the extent they can say: I am. It further follows that such beings have *freedom*, and everything can be imputed to them; and this is *practical personality*, which has consequences in morality. (*ML1*, AA 28:277).[44]

Moral personality is taken to "follow" from psychological personality insofar as the fact that I am conscious of my self as a substance distinct from matter implies that my actions are not determined by external causes.[45] As a result, I must have a capacity for free action, and so my actions can be imputed to me.[46] In contrast to this, then, the mere *psychological* personality at issue for the rational psychologist has no direct "consequences in morality" but amounts simply to the capacity to be conscious of the persistence of our (proper) subject.

If Kant adopts Wolff's conception of the soul's personality, he nonetheless finds reason to doubt Wolff's conclusion that the soul retains its state of personality after death. In the final section of the *ML1* notes, devoted to this topic, Kant takes up two questions essential to the demonstration of the soul's immortality. The first concerns whether the soul will "live and survive" after the death of the body, that is, whether the soul will persist after the death of the body,[47] and whether it will continue its spiri-

---

44 It should be noted that, in the original, the notes refer to "*physikalische*" rather than "*psychologische*" personality. The latter is suggested, however, by the way in which the distinction is presented later in the notes (cf. *ML1*, AA 28:296). For a different version of this distinction, see *R* 5049 (AA 18:72-3).

45 Thus Kant sometimes suggests that practical personality amounts to the consciousness *of* our personality, as in *R* 4228, where he writes that "through consciousness of our personality, we see ourselves in the intellectual world and find ourselves free" (AA 17:467), and *ML1*, AA 28:296, where he distinguishes between the soul's "being conscious to itself that it is a person," which corresponds to practical personality, and its being "conscious of the identity," which corresponds to psychological personality. Contrast Klemme's claim that "[d]ie psychologische Persönlichkeit setzt die praktische Persönlichkeit und damit den Begriff der Freiheit des Subjekts voraus" (1996, 98).

46 See also *R* 4225 (AA 17:464), *R* 4228 (AA 17:467), *R* 5049 (AA 18:72-3), and *R* 6713 (AA 19:139).

47 See *ML1*, AA 28:285: "that [the soul] will merely live already follows from its substantiality, in that every substance survives."

tual existence as a thinking and willing being.[48] Kant claims that this can be proved to hold contingently, but that as regards the "*natural necessity of living*" (*ML1*, AA 28:285) that would be required for immortality, we must be content with arguments that "prove only the *hope* for a future life" (*ML1*, AA 28:285). Like Wolff (and others), however, Kant does not take survival and the continuation of spiritual life to be sufficient for immortality, even claiming that "*personality* [is] the main matter with the soul after death" (*ML1*, AA 28:296), and so it must also be shown that the human soul retains its condition of personality;[49] otherwise, "if it is not conscious of itself, although its life-power is still there, then this is *spiritual slumber*, in which the soul does not know where it is, and cannot rightly adapt itself in the other world" (*ML1*, AA 28:296). It might seem clear that, because the soul's consciousness of itself "rests on inner sense," it should remain a person after death, that is, "because the body is not a principle of life, inner sense [would] remain even without the body, thus likewise the personality" (*ML1*, AA 28:296). Nonetheless, in our case "[e]xperience teaches us that when we think, our body comes into play," even if it is possible for us to imagine other beings for whom this is not the case (*ML1*, AA 28:278). It will be recalled that Wolff had considered the soul's reliance on the organs of sense as a possible obstacle to the demonstration of its immortality, and specifically to its continued spiritual existence. Yet, since Kant understands the very concept we have of the *I*, the basis for the rational psychologist's conception of the soul, to be abstracted from a representation of ourselves that includes the data of outer sense, Kant takes this specter to extend to the soul's personality as well. Consequently, as Kant writes in an earlier note, it is not clear in our case that we could become conscious of ourselves without the involvement of the body:

> One wonders, therefore, whether it is possible that the human soul can be a particular person even without the body [...]. This question does not concern [*enthält*] whether certain interruptions of its personality (as in sleep) take place, *but rather whether it can ever become conscious of itself without a body* (*R* 4237; AA 17:472 – my emphasis).

Kant thus concludes that, if the personality of the soul is to be possible, either it must be the case that (immediately) after death we are provided

48 See *ML1*, AA 28:286, where Kant distinguishes animal from spiritual life, and AA 28:277 where Kant defines spiritual beings as beings "whose thinking and willing can continue even when they are separated from the body."

49 For Kant's recognition of the importance of personality in the demonstrations of the soul's immortality, see *R* 4239 (AA 17:473), *R* 4554 (AA 17:591-2), *R* 5473 (AA 17:192) and *R* 5474 (AA 17:192-3).

with a body, whether "of a worldly or otherworldly kind" (*ML*$_1$, AA 28:296), through which we can receive outer sensations and thereby become conscious of ourselves in the usual way; or, it must be the case that we begin to lead "an entirely pure spiritual life" without a body (*ML*$_1$, AA 28:296) where we become conscious of ourselves through a kind of "*spiritual intuition*" (*ML*$_1$, AA 28:297). While Kant thinks that there is reason to prefer the latter option, he concludes that, as both alternatives far exceed the boundaries of experience (cf. AA 28:264), "we can say nothing reliable" regarding the soul's personality after death (AA 28:296).

3. The Paralogism of Personality

The foregoing context sheds new light on the topic and argumentative strategy of the Third Paralogism. It suggests that, rather than taking up the issue of personal identity, that is, "the question of whether one has an appropriate numerical identity over time," the Third Paralogism will instead concern the issue of personality as traditionally understood, that is, the question of "whether one has certain appropriate complex powers (at any time)."[50] In particular, Kant will reject the rational psychologist's account of *how* we are conscious of our identity, and are thereby persons, claiming that we have no experience of the persistence of the soul to ground this consciousness. Nonetheless, Kant will allow that we are conscious of our numerical identity, in a way distinct from that imagined by the rational psychologist, and which qualifies the soul as a person in the Wolffian sense. This helps to bring out the sometimes overlooked paralogistic character of the rational psychologist's inference to the soul's personality as well as the role of transcendental illusion in motivating this error. Moreover, given that the proper use of the concept of (psychological) personality was not as the ground for the imputation of praise or blame, but rather as a requirement for the demonstration of the soul's immortality, we can make sense of what Kant intends when he claims that the merely transcendental signification of the concept of personality is necessary and sufficient for practical use without resorting to the uncharitable charge of a paralogistic error on Kant's part.

The criticism that follows Kant's reconstruction of the argument for the soul's personality begins by presenting his mature account of the way

50 Ameriks (2000, 129) distinguishes the two questions in this way; nonetheless, he takes the former (the "identity question") to be the sole focus of the Third Paralogism.

in which the soul is "conscious of the numerical identity of its self in different times." Before considering this, it will be useful to look briefly at Kant's previous account in the *ML1* notes of how this consciousness comes about. The relevant discussion is provided in Kant's argument for the soul's *singularity*:

> 3. The soul is a *single soul* (the oneness [*Unität*], the unity [*Einheit*] of the soul), *i.e., my consciousness is the consciousness of a single substance*. I am not conscious of myself as several substances. For if there were several thinking beings in a human being, then one would also have to be conscious of several thinking beings. But the I expresses oneness: I am conscious of myself as one subject. (*ML1*, AA 28:267).

The question at issue in the singularity of the soul naturally arises from the previous demonstrations of the soul's substantiality and simplicity. Having shown that the soul, or *I* (in the narrow sense), "means the subject, so far as it is no predicate of another thing" (*ML1*, AA 28:266), and so is a substance, Kant proceeds to demonstrate that the soul is a simple, rather than composite, substance inasmuch as a "simple thought can take place only in one simple subject" (*ML1*, AA 28:266). Given, then, that the subject of thought must be a simple substance, it remains to be seen whether the *same* (simple) substantial soul is the subject of all our thinking, or whether there are multiple (but non-overlapping) simple subjects.[51] As Kant argues here, the singularity of my soul is evidenced by my consciousness of myself as a single substance. Upon attending to myself, I notice that the *I*, or soul, persists through the succession of my thoughts, and so is the same substance throughout. According to Kant, were it the case that the soul was in fact comprised of consecutive substances it would, upon directing its attention to itself, notice this and its consciousness would be of itself as "several substances." Thus, the soul is singular inasmuch as its consciousness of itself as single or a unity could only be grounded in the sameness of its substance.

By the time of the *KrV*, of course, Kant had come to think differently of how we are conscious of the numerical identity of ourselves in different times, and his new account is presented in the first paragraph of the Criticism that follows the reconstructed argument. There, Kant briefly outlines

---

51 Kant is clear that the singularity of the soul does not merely amount to its simplicity: "die Unität ist mit der Simplicität nicht einerlei" (*ML1*, AA 28:263). Moreover, as it is the consciousness of the soul's singularity, or unity (as distinct from simplicity) that is at issue in the Third Paralogism, it should be clear that the argument does not collapse into the Second Paralogism. Contrast Ameriks (2000, 129).

the way in which we cognize (and so, become conscious) of the numerical identity of an external object:

> If I want to cognize through experience the numerical identity of an external object, then I will attend to what is persisting in its appearance, to which, as subject, everything else relates as a determination, and I will notice the identity of the former in the time in which the latter changes. (*CPR* A 361-2).

As should be clear, the way in which Kant now claims that we cognize the numerical identity of an *external* object is precisely the way in which he had claimed that the singularity of the soul was cognized in the ML_1 notes. As Kant goes on to elaborate, however, he now holds that we become conscious of the identity or unity of the soul in a much different way. Far from having to attend to what is persistent in the appearance of this unitary consciousness in order to cognize its identity, the soul, or the *I* of the *I think*, is necessarily represented as numerically identical. This is because the possibility of the cognition of objects requires the capacity to combine my representations, which itself presupposes that all my representations belong to a numerically identical consciousness.[52] That the *consciousness* of this numerical identity of apperception,[53] which is possible with respect to all my representations (cf. *CPR* B 408), is furthermore a consciousness of the numerical identity of myself *in different times* is evident upon a consideration of the transcendental ideality of time:

> But now I am an object of inner sense and all time is merely the form of inner sense. Consequently, I relate each and every one of my successive determinations to the numerically identical self in all time, i.e., in the form of the inner intuition of my self. (*CPR* A 362).

Time is a form of intuition which is also to say that it is a representation (a pure intuition), and as such its unity must be taken to originate in the unity of apperception.[54] This implies that *all* determinations given in time are capable of being related to the numerically identical self, and so that

---

52 Kant refers to apperception as *numerically* identical at *CPR* A 107 ("That which should *necessarily* be represented as numerically identical") and A 114 ("But from this, as a transcendental representation, numerical identity is inseparable").

53 See *CPR* A 116: "We are conscious *a priori* of the thoroughgoing identity of ourselves with regard to all representations that can ever belong to our cognition, as a necessary condition of the possibility of all representations (since the latter represent something in me only insofar as they belong with all the others to one consciousness, hence they must at least be capable of being connected in it)."

54 As Kant had written previously, "even the purest objective unity, namely that of *a priori* concepts (space and time) is possible only through the relation of intuitions to it [i. e., *transcendental apperception*]. The numerical unity of this apperception therefore grounds all concepts *a priori*" (*CPR* A 107).

the self of whose identity I am conscious is necessarily identical in all the time in which I am possibly conscious.

The rational psychologist (including Kant at the time of the *ML1* lectures) is thus accused of conflating the way in which we are conscious of the numerical identity of the *I* of the *I think* with the way in which we are conscious of the numerical identity of external objects. Significantly, this diagnosis of the rational psychologist's error is quite consistent with Kant's general account of the error in each paralogism as involving a middle term ambiguous between transcendental and empirical uses.[55] Kant distils the rational psychologist's argument for the personality of the self into the following syllogism (with implied information supplied):

> What is conscious of the numerical identity of itself in different times is to that extent a *person*.
> Now the soul [or thinking *I*] is [conscious of the numerical identity of its self in different times].
> Thus it is a person. (*CPR* A 361).

At the end of the Paralogisms chapter, Kant provides the following general account of the formal error committed in each of the four paralogisms:

> If one wants to give a logical title to the paralogism in the dialectical syllogism of the rational doctrine of the soul, insofar as they have correct premises, then it can count as a *sophisma figurae dictionis*, in which the major premise makes a merely transcendental use of the category, in regard to its condition, but in which the minor premise and the conclusion, in respect of the soul that is subsumed under this condition, make an empirical use of the same category. (*CPR* A 402-3).

As Kant explains later in his *Logic*, the fallacy at issue in a paralogism is that of an ambiguous middle term. In the case of the third paralogism, the middle term is "that which is conscious of the numerical identity of itself in different times" which, in referring to *identity* evidently makes use of the category of unity.[56] As it occurs in the middle term of the major premise (which reproduces the Wolffian definition of person[57]), the category is made use of unambiguously in its transcendental signification, that is, it serves only to specify what condition an object in general must meet in order to be subsumed under that concept. It is in the minor premise that the problem occurs, and the ambiguity in the case of the third paralogism involves the concept of numerical identity, which Kant takes to be employed in both its transcendental and empirical uses in that premise. In its

---

55 Contrast Ameriks (2000, 130).

56 Thus Kant uses 'unity' interchangeably with 'identity' at *CPR* A 107 ("The numerical unity of this apperception"), and the category of unity is explicitly linked to identity in the charts at A 344/B 402 and A 404.

57 Klemme (1996, 329) also notes this.

empirical signification, the concept of unity refers to that which is numerically one in a given time (and space),[58] and for something to be numerically identical *in different times*, any changes to it over time must merely be changes of its state, through which it persists.[59] This is in contrast to its transcendental signification, that is, its sense in its application to objects in general and, specifically, to the *I* of the *I think*, where the identity of such an object at different times consists in its logically necessary unity rather than in its persistence over time. Kant contrasts these two senses of identity briefly at *CPR* A 365 ("Now since *this* identity of person in no way follows from the identity of the I") and in more detail in the second edition version:

> The proposition of the identity of myself in everything manifold of which I am conscious is equally one lying in the concepts themselves, and hence an analytic proposition; but *this* identity of the subject, of which I can become conscious in every representation, does not concern the intuition of it, through which it is given as an object, and thus cannot signify the identity of the person, by which is understood the consciousness of the identity of its own substance as a thinking being in all changes of state (*CPR* B 408 – my emphasis).

Accordingly, the rational psychologist is guilty of conflating these two senses of the middle term in order to run his argument. This generates a syllogism with four terms, a paralogism, and leads the rational psychologist to commit a further, metaphysical error, a "subreption of hypostatized consciousness" (*CPR* A 402), as the rational psychologist predicates a category as it is used with regard to objects of our experience or empirically of the soul, which is not given as an object of sensible intuition.

The conflation of the empirical cognition of an identical object persisting over time with a mere formal consciousness of unity might seem an unlikely error, even for the rational psychologist. According to Kant, however, it has its roots in a natural and unavoidable illusion which he generally defines as the "taking of a *subjective* condition of thinking for the cognition of an *object*" (*CPR* A 396).[60] In the case of the idea of the *I* of the *I think*, or soul, it is the representation of the unconditioned unity of inner appearances and, as such, is only presumed as given in the course of reason's own efforts to bring unity into that manifold; yet it inevitably seems

---

58 See *CPR* A 263/B 319-20.

59 Thus we find in the First Analogy: "Alteration is a way of existing that succeeds another way of existing in the very same object. Hence everything that is altered is *lasting*, and only its *state changes*" (*CPR* A 187/B 230).

60 For more detail on the distinction between transcendental illusion and the paralogistic error it grounds, see Grier (2001, 144-52) and Allison (2004, 334-42).

to be given as an object in its own right. Thus, Kant writes near the conclusion of the Paralogisms that "nothing is more natural and seductive than the illusion of taking the unity in the synthesis of thoughts for a perceived unity in the subject of these thoughts" (*CPR* A 402).[61] As it relates to the topic of the third paralogism, the rational psychologist is misled by the illusory representation of the soul into taking the merely formal unity of apperception for a given unity and, on that basis, mistakes our consciousness of our numerical identity in different times for the consciousness of the identity of our own immaterial substance. Consequently, the rational psychologist illicitly takes the personality of the soul to be founded on the cognition of the identity in different times of an object given in time, when in fact all that we can have is a mere consciousness of the unity of the logical subject of thought.

Indeed, the error the rational psychologist commits on the basis of this illusion is shown to be all the more egregious given that a consciousness of the self as a numerically identical persisting being in time is only possible from the perspective of an external observer, and even then does not amount to a consciousness of the identity of my (proper) self:

> But if I consider myself from the standpoint of another (as an object of his outer intuition), then it is this external observer who originally considers *me* as *in time*; for in apperception *time* is properly represented only *in me*. Thus from the I that accompanies – and indeed with complete identity – all representations at every time in *my* consciousness, although he admits this I, he will still not infer the objective persistence of my self. (*CPR* A 362-3).

Against these charges, the rational psychologist might reply along the lines of Kant's previous exposition in the *ML1* notes that my own consciousness of the numerical identity of the *I* of the *I think* proves the identity of my substance since, if I was comprised of consecutive substances, I could not be conscious of my numerical identity. Indeed, this is the very scenario that Kant now lays out in detail in the section's notorious footnote:

> An elastic ball that strikes another one in a straight line communicates to the latter its whole motion, hence its whole state [...]. Now assuming substances, on the analogy with such bodies, in which representations, together with consciousness of them, flow from one to another, a whole series of these substances may be thought, of which the first would communicate its state, together with its consciousness, to the second, which would communicate its own state, together with that of the previous substances, to a third substance, and this in turn would share the states of all previous ones, together with their consciousness and its own. (*CPR* A 363-4n).

61 Compare *R* 5553 (AA 18:223).

As should be clear, the significance of this note lies primarily in the fact that Kant now takes issue with his previous claim in the ML_1 notes that, "if there were several thinking beings in a human being, then one would also have to be conscious of several thinking beings" (28:267). In the fabricated scenario, Kant imagines that a human being is composed of several consecutive, non-overlapping thinking beings, in which thoughts and their accompanying consciousness are transferred from one thinking substance to the next, in a manner analogous with the transfer of motion among corporeal substances. Since for Kant, as for Wolff, the identity of person is just the identity of underlying substance, the assumption of multiple substances underlying our states of thought and consciousness obviously implies that the human being "would not have been the very same person in all these states" (*CPR* A 364n). Even so, Kant now contends that the consciousness to which our thoughts belong, as designated by the *I* and which is transferred along with those thoughts from substance to substance, would necessarily remain numerically identical throughout, despite its inherence in a number of substances; as he writes "a [substantial] change can go on that does not allow it [i.e., my subject] to keep its identity; and this even though all the while the consonant [*gleichlautende*] I is assigned to it" (*CPR* A 363). Accordingly, without pronouncing on the question of whether we would be conscious of the fact if we *were* comprised of multiple non-overlapping thinking substances, Kant claims that a consciousness of the numerical identity of the *I* of the *I think* would still be possible since its identity is in no way tied to the identity of an underlying thinking substance.

Significantly, Kant continues to hold that, even on the basis of its consciousness of its numerical identity, where this identity is taken in its transcendental signification, the soul is rightly taken to be a *person*. As Kant writes:

> On this basis, the personality of the soul must be regarded not as inferred but rather as a completely identical [i. e., analytic[62]] proposition of self-consciousness in time, and that is also the cause of its being valid *a priori*. For it really says no more than that in the whole time in which I am conscious of myself, I am conscious of this time as belonging to the unity of my self, and it is all the same whether I say that this whole time is in Me, as an individual unity, or that I am to be found with numerical identity in all this time. (*CPR* A 362).

Indeed, it is with respect to personality so understood that Kant makes the troublesome claim that "to this extent this concept is also necessary and sufficient for practical use" (*CPR* A 365-6). Though it had previously been

62 See Ameriks (2000, 131).

assumed that the practical use which Kant imagines for the concept here could only be that forensic use emphasized by Locke, which assumption ultimately issues in Longuenesse's charge of a paralogism of pure practical reason, it should now be clear that this need not be the case. Instead, the "practical use" to which the concept of personality is put could only be that to which Wolff, and those he influenced, put it, namely, its use in the proof of the immortality of the soul. Along these lines, that the personality of the soul, understood in this sense, is necessary and sufficient for this use should be taken to mean that, provided the soul survives death (about which we cannot be certain since we do not cognize its substantiality and simplicity), it will remain conscious of its numerical identity, and thus will retain its *status personalitatis*. That Kant should now make this claim represents another clear departure from the ML_1 notes where Kant had concluded that nothing reliable could be known about the personality of the soul after death on the grounds that our experience was of the necessity of outer intuition, and therefore of the body, for such consciousness. The likely reason for this departure is that Kant no longer holds that the concept of the soul at issue in rational psychology, and which expresses the consciousness of our own identity, is drawn from a prior representation that includes the data of outer sense. Rather, as he now claims, the concept of the *I* is inferred (by means of reason) from "the transcendental concept of a subject that contains nothing manifold" (*CPR* A 340/B 397-8), and consequently, that the "numerical identity of our self," of the *I* of the *I think*, is concluded "from identical apperception" (*CPR* A 365). Given that nothing else is required for the soul to retain its state of personality after death than the possible consciousness of the numerical identity of apperception, the personality of the soul is easily conceivable even in the absence of the body and all the data of outer sense. This does not, of course, imply that we can be certain of the soul's survival of the death of the body as a pure spiritual being since, as we have seen, the identity of the *I* of the *I think* does not prove the identity of the thinking substance in which it inheres; so, even while we can be assured that the surviving soul will retain the state of personality after death "we can never boast of it as an extension of our self-knowledge through pure reason" (*CPR* A 366).

Still, it might be wondered whether Kant's understanding of the transcendental ground for the personality of the soul does not deviate from the spirit of the traditional concept. The initial worry on the part of the rational psychologist, as was reflected in the appropriation of the classical myth of a draught of forgetfulness, was that the soul would not retain a memory of itself, of its actions and experiences in this life, which circum-

stance would give us reason to doubt the wisdom of this arrangement of the world. Admittedly, Kant's assertion that the soul would retain its state of personality, where this consists merely in the consciousness of the "logical identity of the I" (*CPR* A 363), can offer little consolation in this regard. Nonetheless, we have seen that a part of the rational psychologist's worry had to do with whether the human soul will retain its distinctive capacities after death. When taken in this sense, a draught of forgetfulness would threaten the loss of those capacities whose presence was signalled by the soul's personality and which served to distinguish the human soul from that of the animal. Understood in this context, Kant's argument performs much better. For Kant, the soul's consciousness of its numerical identity, as ultimately a consciousness of the identity of *apperception*, is evidence of the soul's possession of an understanding.[63] Animals, by contrast, do not possess an understanding but only its analogue; as Kant writes in the *Anthropology* "irrational animals [*verstandlosen Thiere*] can manage provisionally, following implanted instincts, like a people without a sovereign" (AA 7:196; § 40). That the personality of the human soul signals the possession of an understanding thus serves to distinguish the human soul from that of an animal, something that Kant also makes clear in the *Anthropology*:

> That man can have the I among his representations elevates him infinitely above all other living beings on earth. He is thereby a person [...] that is, by rank and worth a completely distinct being from things that are the same as reason-less animals with which one can do as one pleases. (AA 7:127; § 1).[64]

Accordingly, even when the personality of the soul is taken in its transcendental signification, it suffices to distinguish us, as beings with an understanding, from "*die verstandlosen Thiere*." Rather than betraying a paralogism of pure practical reason, then, Kant is entirely warranted in his claim that the soul's personality, its consciousness of the numerical identity of the *I* of the *I think* in all time, suffices for its practical use. Since the personality of the soul depends in no way upon the existence of the body, Kant can claim to have shown that the human soul would retain its distinctive capacities in the afterlife.[65]

---

63 See *CPR* A 119: "*The unity of apperception in relation to the synthesis of the imagination* is the *understanding*" and B 134n: "the synthetic unity of apperception is the highest point to which one must affix all use of the understanding [...] indeed, this faculty is the understanding itself."

64 Cf. also the letter to Herz of May 26th, 1789 (AA 11:52).

65 I am grateful to Falk Wunderlich, Lorne Falkenstein, Ben Hill, an anonymous referee for the *Kant Yearbook*, and to Heiner Klemme and the participants in his

Bibliography

Allison, Henry E. (1977): Locke's Theory of Personal Identity: A Re-Examination, in: I. C. Tipton (ed.): Locke on Human Understanding: Selected Essays, Oxford, pp. 105-22.

Allison, Henry E. (2004): Kant's Transcendental Idealism: An Interpretation and Defense, revised and enlarged edition, New Haven.

Ameriks, Karl (2000): Kant's Theory of Mind: An Analysis of the Paralogisms of Pure Reason, new edition, Oxford.

Beck, Lewis White (1969): Early German Philosophy, Cambridge MA.

Calvin, John (1932): Psychopannychia, ed. W. Zimmerli, Leipzig.

Canz, Israel Gottlieb (1744): Uberzeugender Beweis aus der Vernunft von der Unsterblichkeit sowohl der Menschen Seele insgemein, als besonders der Kinder-Seelen, Tübingen.

Descartes, René (1999): The Philosophical Writings of Descartes, ed. and trans. J. Cottingham, R. Stoothoff, and D. Murdoch, Cambridge.

Dyck, Corey W. (2009): The Divorce of Reason and Experience: Kant's Paralogisms of Pure Reason in Context, in: Journal of the History of Philosophy 47, pp. 249-75.

Grier, Michelle (2001): Kant's Doctrine of Transcendental Illusion, Cambridge.

Kant, Immanuel (1900 ff.): Gesammelte Schriften, ed. königlich preußische (later German) Akademie der Wissenschaften, Berlin.

Kant, Immanuel (1998): Critique of Pure Reason, ed. and trans. P. Guyer and A. W. Wood, Cambridge.

Kant, Immanuel (2001): Lectures on Metaphysics, ed. and trans. K. Ameriks and S. Naragon, Cambridge.

Kant, Immanuel (2007): Anthropology, History, and Education, ed. G. Zöller and R. Louden, Cambridge.

Klemme, Heiner (1996): Kants Philosophie des Subjekts, Hamburg.

Leibniz, Gottfried Wilhelm (1990): Theodicy, ed. A. Farrer, trans. E. M. Huggard, Chicago.

Locke, John (1979): An Essay Concerning Human Understanding, ed. P. H. Nidditch, Oxford.

Longuenesse, Béatrice (2007): Kant on the Identity of Persons in: Proceedings of the Aristotelian Society, pp. 149-67.

Meier, Georg Friedrich (1749): Gedancken von dem Zustande der Seele nach dem Tode, second edition, Halle.

Mendelssohn, Moses (1979): Phaedon oder über die Unsterblichkeit der Seele, Hamburg.

Oberseminar at the Johannes Gutenberg-Universität Mainz for helpful comments on earlier drafts of this paper.

Reinbeck, Johann Gustav (1739): Philosophische Gedancken über die vernünfftige Seele und derselben Unsterblichkeit, Berlin.

Thümmig, Ludwig Philipp (1721): Demonstratio immortalitatis animae ex intima ejus natura deducta, Halle.

Thümmig, Ludwig Philipp (1725): Institutiones philosophiae Wolfianae, Frankfurt and Leipzig.

Virgil (1997): Virgil's Aeneid, trans. J. Dryden, New York.

Wolff, Christian (1733): Ausführliche Nachricht von seinen eigenen Schrifften, die er in deutscher Sprache von den verschiedenen Theilen der Welt-Weisheit heraus gegeben, Frankfurt.

Wolff, Christian (1736): Prima philosophia sive Ontologia, second edition, Frankfurt, Leipzig.

Wolff, Christian (1738): Psychologia empirica, second edition, Frankfurt, Leipzig.

Wolff, Christian (1740): Vernünftige Gedancken von Gott, der Welt, und der Seele des Menschen, Anderer Theil, bestehend in ausführlichen Anmerckungen [Anmerkungen zur Deutschen Metaphysik], fourth edition, Frankfurt.

Wolff, Christian (1740-8): Jus naturae, 8 vols., Frankfurt, Leipzig.

Wolff, Christian (1740): Psychologia rationalis, second edition, Frankfurt, Leipzig.

Wolff, Christian (1740): Philosophia rationalis, sive Logica, 3 vols., third edition, Frankfurt, Leipzig.

Wolff, Christian (1751): Vernünfftige Gedancken von Gott, der Welt und der Seele des Menschen [Deutsche Metaphysik], eleventh edition, Halle.

Wundt, Max (1945): Die deutsche Schulphilosophie im Zeitalter der Aufklärung, Tübingen.