

Violi Sahaj, Bureau 120.049, Blandijnberg 2, 9000-Gent, Belgium, e-mail: violi.v.sahaj@gmail.com; Prof. Dr. Bart Vandenabeele, Bureau 110.072, Blandijnberg 2, 9000-Gent, Belgium, e-mail: Bart.Vandenabeele@ugent.be

Violi Sahaj / Bart Vandenabeele

Nietzsche and the Limitations of the Transcendental Reading

Abstract: The Nietzschean conception of selfhood has been the subject of considerable debate in the Anglophone commentary. This debate has been focused on what Sebastian Gardner coined as ‘the lack of fit’ between Nietzsche’s theoretical and practical remarks on the self. There have been various attempts at a solution to the lack of fit and in this article we address one such solution, which we call the ‘transcendental reading’. We argue that the reading is right to highlight that Nietzschean selfhood risks elimination of first-person practical agency. **We contend, however, that the reading limits our understanding of his critique of a strictly first-person conception of selfhood. This critique aims to reject a conception of the self as distinct from the drives.** We finally suggest an alternative solution to the lack of fit that takes into account the concerns of the transcendental reading, but seeks to overcome its limitations.

Keywords: Consciousness, Self-conscious ‘I’, Subjectivity, Selfhood, Agency, Eliminativism, First-person Perspective, Third-person Perspective

Zusammenfassung:

Schlagwörter: Das Bewusstsein, selbstbewusst ich, Subjektivität, Selbstheit, Agentur, Eliminativismus, Ich-Perspektive, Dritte-Person Perspektive

The diverse claims Nietzsche makes on selfhood have been the subject of considerable debate centred on the conception of selfhood to which he commits himself, or towards which he aims to lead his readers. These claims are scattered throughout his works, but at least some of them can be grouped into three claims that suggest a candidate for a Nietzschean picture on the self. First, he rejects the existence of such a ‘thing’ as the ‘will’, ‘self’, ‘ego’ or the ‘I’, which he construes as resting on belief in the ‘causality of the will’:

The ‘inner world’ is full of illusions and phantasms: will is one of them. The will does not do anything anymore, and so it does not explain anything anymore either — it just accompanies processes, but it can be absent as well. The so-called ‘motive’: another error. Just a surface phenomenon of consciousness,

an ‘after-the-fact’ that hides the *antecedentia* of an act more than it reveals them. Not to mention the I [*das Ich*]! That has become a fairy tale, a fiction, a play on words: it has stopped thinking, feeling, and willing altogether! . . . What follows from this? There are no mental [*geistigen*] causes whatsoever! (TI, ‘The Four Great Errors’, 3)

Secondly, and somewhat in opposition to the above, there are passages such as the following where Nietzsche draws a distinction between what he calls the ‘real ego’ and the ‘phantom of the ego’:

Whatever they may think and say about their ‘egoism’, the great majority nonetheless do nothing for their ego their whole life long: what they do is done for the phantom of their ego which has formed itself in the heads of those around them and has been communicated to them; — as a consequence they all of them dwell in a fog of impersonal, semi-personal opinions, and arbitrary, as it were poetical evaluations, the one for ever in the head of someone else, and the head of this someone else again in the heads of others: a strange world of phantasms — which at the same time knows how to put on so sober an appearance! (D 105)²

The majority of us are committed to acting for the ‘phantom of our ego’, but Nietzsche does not stop there. Further in the same passage he defends the possibility of a ‘real’ ego not accessible to an individual in this ‘great majority’; the individual to which he is referring is presumably the ‘sovereign individual’.³ He then proceeds to elaborate on the distinction between the ‘real’ and the ‘phantom’ of our ego in the following way:

¹ A little earlier in the same passage Nietzsche claims that concepts such as ‘thing’, ‘self’, ‘ego’ or the ‘I’ rest on the belief that the will is causally efficacious: “Of all these three ‘inner facts’ that together seem to guarantee causation, the first and most convincing is that of *will as causal agent*; the conception of a consciousness (‘mind’) as cause, and then that of the I (the ‘subject’) as cause are just latecomers that appeared once causality of the will was established as given, as empirical . . . Nowadays we do not believe a word of it.” (TI, ‘The Four Great Errors’, 3) We quote Friedrich Nietzsche, *The Anti-Christ, Ecce Homo, Twilight of the Idols (A/EC/TI)*, Aaron Ridley and Judith Norman (Eds.), Judith Norman (Trans.), Cambridge 2005.

² We quote Friedrich Nietzsche, *Daybreak: Thoughts on The Prejudices of Morality (D)*, Maudemarie Clark and Brian Leiter (Eds.), R.J. Hollingdale (Trans.), Cambridge 1997, p. 61.

³ Nietzsche uses Goethe as an example of the individual who sets up his own ego which annihilates the ‘general pale fiction’ of his age: “In the middle of an age inclined to unreality, Goethe was a convinced realist: he said yes to everything related to him, — his greatest experience was of that *ens realissimum* that went by the name of Napoleon.” (TI, *Skirmishes of an Untimely Man*, 49).

— all these people, unknown to themselves, believe in the bloodless abstraction ‘man’, that is to say, in a fiction; and every alteration effected to this abstraction by the judgments of individual powerful figures (such as princes and philosophers) produces an extraordinary and grossly disproportionate effect on the great majority — all because no individual among this majority is capable of setting up a real ego, accessible to him and fathomed by him, in opposition to the general pale fiction and thereby annihilating it. (D 105)⁴

Thirdly, and most importantly for the current debate, there are passages where he argues that we cannot possibly act without some reference to ourselves and so without a self to which our actions refer:

No man has ever done anything that was done wholly for others and with no personal motivation whatever; how, indeed, should a man be *able* to do something that had no reference to himself, that is to say lacked all inner compulsion (which would have its basis in a personal need)? How could the ego act without the ego? (HH I: 133)

We recognize three claims about the self that do not seem to hang together in such a way as to offer us a Nietzschean picture of selfhood. First, he rejects the existence of the ‘I’, though we should add that what he rejects is its conception as a ‘thing’, which implies an epiphenomenalist or eliminativist stance on the self as reducible to material causes or relations.⁵ Second, he rejects our self-knowledge claims by arguing that we are prone to confuse ‘our ego’ (the real ego) with other people’s opinions of us (the ‘phantom of our ego’), which suggests that he has a normative account of genuine selfhood. Third, he claims that we

⁴ For a better understanding of what Nietzsche is aiming for, it is useful to compare this passage to the following, “*Will a self*”. - Active, successful natures act, not according to the dictum ‘know thyself’, but as if there hovered before them the commandment: *will* a self and thou shalt *become* a self. - Fate seems to have left the choice still up to them; whereas the inactive and contemplative cogitate on what they *have* already chosen, on *one* occasion, when they entered into life” (HH II: 366) we quote Friedrich Nietzsche, *Human all too Human* (HH I & II / WS), R.J. Hollingdale (Trans.), Cambridge 1996.

⁵ We can juxtapose this suggestion to the following passage that openly rejects materialism and voices his own aims, “While Copernicus convinced us to believe, contrary to all our senses, that the earth does *not* stand still, Boscovich taught us to renounce belief in the last bit of earth that *did* “stand still,” the belief in “matter,” in the “material,” in the residual piece of earth and clump of an atom: it was the greatest triumph over the senses that the world had ever known. – But we must go further still and declare war – a ruthless fight to the finish – on the “atomistic need” that, like the more famous “metaphysical need,” still leads a dangerous afterlife in regions where nobody would think to look” (BGE 12) we quote Friedrich Nietzsche, *Beyond Good and Evil* (BGE), Rolf-Peter Horstmann / Judith Norman (Eds.), Judith Norman (Trans.), Cambridge 2002.

cannot possibly act without some reference to ourselves, which he here construes as acting from some ‘inner compulsion’. Attempts to reconcile these diverse and ostensibly conflicting claims on selfhood have instigated a debate in the literature centred on what has come to be known as the ‘lack of fit’ argument.

There are some attempts to resolve the lack of fit that, we claim, have their limitations and fall short of providing a satisfactory solution. In this paper we assess *one* of these attempts and its limitations, which we have coined as the ‘transcendental reading’. Some authors have suggested that Nietzsche must commit himself to a transcendental stance on agency by construing agential actions as grounded on the self-conscious ‘I’. This reading argues that Nietzsche should adopt a strategy similar to Kant, who in the second edition of the Critique of Pure Reason famously claimed that,

if the critique has not erred in teaching that the object should be taken in a *twofold meaning*, namely as appearance or as thing in itself; if its deduction of the pure concepts of the understanding is correct, and hence the principle of causality applies only to things taken in the first sense, namely insofar as they are objects of experience, while things in the second meaning are not subject to it; then just the same will be thought of in the appearance (in visible actions) as necessarily subject to the law of nature and to this extent not free, while yet on the other hand it is thought of as belonging to a thing in itself as not subject to that and hence free, without any contradiction hereby occurring.⁶

For Kant, we can have a transcendental conception of ourselves as free and independent from our experience. When we try to make ourselves an object of experience, however, we do not conceive ourselves as free and independent, but as subjected to the same physical laws as any other object of experience. We can do the previous, Kant believes, without contradiction. We notice in Nietzsche, however, a vociferous rejection of the legitimacy of our positing things in themselves *wholly* independent from experience. Nietzsche’s concern is not with the posit of a thing in itself as a limiting concept whose relationship to experience is negative and therefore empty of significance⁷, but that this concept acquires significance through the moral actions it is supposed to underpin or explain.

⁶ Immanuel Kant, Critique of Pure Reason, Paul Guyer / Allen W. Wood (Eds. & Trans.), New York 1998, p.116 (B xxvii-xxviii).

⁷ See also, “Perhaps we shall then recognize that the thing in itself is worthy of Homeric laughter: that it appeared to be so much, indeed everything, and is actually empty, that is to say empty of significance” (HH I: 16)

“Your insight into *how such things as moral judgements could ever have come into existence* would spoil these emotional words for you, as other emotional words, for example, ‘sin’, ‘salvation of the soul’, and ‘redemption’ have been spoiled for you. And now don’t bring up the categorical imperative, my friend! The term tickles my ear and makes me laugh despite your very serious presence. I am reminded of old Kant, who helped himself to (*erschlichen*) the ‘thing in itself’ — another very ridiculous thing! — and was punished for this when the ‘categorical imperative’ crept into (*beschlichen*) his heart and made him stray back to ‘God’, ‘soul’, ‘freedom’, ‘immortality’, like a fox who strays back into his cage.” (GS 335)⁸

We should consequently be careful not to confuse Nietzsche’s transcendental stance on the self, if he had one, with acceptance of the Kantian conception of agency, which for Nietzsche includes the moral consequences of that conception.⁹ Whether Nietzsche is permitted to have a transcendental conception of agency without this conception being Kantian is, we believe, a question beyond the scope and consideration of the current paper.

In this paper, we will discuss two important transcendental readings of Nietzsche, one by Sebastian Gardner (in section 2) and one by Chris Janaway (in section 3). Both Gardner’s and Janaway’s readings of Nietzschean agency are right to argue that we cannot eliminate the ‘I’ without also eliminating first-person practical agency. They are equally right to argue that there is a lack of fit in our best interpretations of Nietzsche, if not in his own thoughts. We argue (in section 4) that Nietzsche had an alternative conception of agency that debunks the causal role of the ‘I’, but as suggested by the transcendental reading, this conception should not eliminate first-person practical agency. We submit, however, that construing his critique of the causal efficacy of the ‘I’ to imply its elimination is wrong. Instead, we claim (in section 5) that Nietzsche’s critique is better construed as putting into question the supposed influence the ‘I’ has over our actions. This critique is aimed at the relation between the ‘I’-utterance and

⁸ We quote Friedrich Nietzsche, *The Gay Science* (GS), Bernard Williams (Ed.), Josephine Nauckhoff and Adrian del Caro (Trans.), Cambridge 2001.

⁹ For a detailed and useful analysis of Nietzsche’s relationship to Kant see R. Kevin Hill, *Nietzsche’s Critiques: The Kantian Foundations of his Thought*, New York 2003, who is right to emphasise the differences between the two thinkers on agency: “whereas Kant believes that behind the merely synthetic unity of the apperceiving self, there is a noumenal self, Nietzsche claims that there is no such thing. There is only the body. A genuine substance must not be adjectival on any other substance or attribute and it must not be composed of parts. The body is composed of parts. Thus there is no substantial self, either ‘here’ or ‘elsewhere’. This, Nietzsche takes it, represents a critical advance beyond Kant’s position, and depends crucially upon his rejection of things-in-themselves.” (p. 181) For further and a more in depth reading on Kant’s conception of selfhood see Henry E. Allison, *Kant’s Transcendental Idealism: An Interpretation and Defense*, New Haven 2004, pp. 333-356.

its corresponding actions, which, we argue, reveals what we are calling the bridge between the first-person and third-person conceptions of selfhood. Nietzsche therefore seeks to argue for a different, more original and perhaps normative account of what constitutes 'genuine' agency. Towards the end of the paper we suggest an alternative solution and approach to the lack of fit problem.

1. The Self and the 'Lack of Fit' Argument

The 'lack of fit' is coined by Gardner¹⁰, whose reading of Nietzschean selfhood argues that there is a lack of fit between Nietzsche's practical and theoretical claims on the self and that his philosophical project sought to reconcile them. The lack of fit then claims that Nietzsche's practical account construes agency as a process of self-valuation, self-determination and self-affirmation illustrated by his views on the sovereign individual. His theoretical account, however, construes the 'I' as a fiction and-or as reducible to hierarchal power relations between drives.

Nietzsche describes the I as an illusion and a fiction ... it is not just the I of the philosophers, or the I as it may appear in reflection, but the I in all respects and contexts of consideration that is held to be illusory.¹¹

But, Gardner argues, in order for the sovereign individual to engage in the process of self-valuation and so on, she must first believe in the reality of the 'I' which she uses in speaking and thinking about herself. She cannot do this, however, if this 'I' is construed as a fiction or if it is reduced to power relations between drives. Accordingly, the self in practice and the self in theory do not meet on a common ground.

Nietzsche's theoretical account does not offer enough argument for reducing the unity of the 'I' to power relations between drives, nor for explaining how (the belief in) this fiction arose in the first place. The first-person perspective of the 'I' does not meet with the third-person perspective of what the self is construed as. If the 'I' in general is illusory, then how do we account for the sovereign individual's practical agency, which requires the belief in the 'I'? Further, and granting that the sovereign individual practises agency on the basis of an

¹⁰ Sebastian Gardner, Nietzsche, the Self, and the Disunity of Philosophical Reason, in: Ken Gemes / Simon May (Eds.), Nietzsche on Freedom and Autonomy, New York 2009, pp. 1-31.

¹¹ Gardner, Nietzsche, the Self and Disunity, p. 2.

illusion, how do we explain her belief in this illusion? This, in a nutshell, is what the ‘lack of fit’ amounts to and the problems that face Nietzsche’s account of the self and agency.

Gardner proposes a solution to the lack of fit presumably without loss to Nietzsche’s central aims. The self-conscious unity of the ‘I’ has a transcendental, not a “transcendentally realistic” status.¹² This allows that Nietzsche preserve his objections to the ‘I’, and so is sensitive to his theoretical claims, but does not lose the reality of the ‘I’, which is inherent to practical agency. The sovereign individual does not require the ‘I’ to have the status of “a substratum, nor as equipped with freedom in the indeterministic sense”, but “as occupying the position of ground”.¹³ This solution must hold “together in a coherent manner, *both* the unitary I of self-consciousness *and* the psychological manifold”.¹⁴

One might clarify Gardner’s proposal via a distinction between cause and ground.¹⁵ The cause of the content of the I-thought can be explained by interactions between third-person composites (ibid.). For example, the thought “I want X” can be explained in terms of ‘X’ being the effect of an interaction between drives that form hierarchal power relations. The unity of the ‘I’ that ‘wants X’, however, is independent from and serves as the ground for the composites and their hierarchy. Within this account is also the modest contention that this

12 Gardner illustrates this by arguing that, “there is a boundary to be marked and respected between legitimate transcendental claims for the necessity of the I-representation, and illegitimate, transcendently realistic claims regarding the constitution-in-itself of a corresponding object of this representation” Gardner, Nietzsche, the Self and Disunity, p. 12. Presumably, in denying the reality of the ‘I’, Nietzsche makes this illegitimate claim.

13 Gardner, Nietzsche, the Self and Disunity, p. 9.

¹⁴ Gardner, Nietzsche, the Self and Disunity, p. 13.

15 Gardner describes the solution in the following manner, “this would allow it to be maintained **both** that the causality of the I is properly interpreted as involving necessarily the expression of the dominant power-unit in the psychological composite (or whatever functional arrangement Nietzschean theory tells us is involved), **and** that the thought, which the Nietzschean subject must entertain when a power-unit realizes itself successfully, wills values, etc., is an **I-thought**, not the thought **that** such and such a power-unit or whatever prevails presently” Gardner, Nietzsche, the Self and Disunity, p. 13. Nietzsche advocates a similar distinction between cause and ground, **but contends that its utility and indispensability “proves nothing against its fictionality: a belief can be a condition of life and still be false [beweist noch nichts gegen ihre Erdichtetheit: es kann ein Glaube Lebensbedingung und trotzdem falsch sein]”** (Nachlass, 38[3], KSA 11.598; our translation). Thus he is aware of the implications of abandoning the ‘I’ (and not simply its causality), as he says in **Nachlass 7[63], KSA 12.317 (our translation): “to let it go means to no longer be able to think [ihn fahren lassen heißt nicht-mehr-denken-dürfen]”**. Is this abandonment what Nietzsche is aiming for? It is our aim to argue that it is not.

self-conscious unity is presupposed by every act of thought and valuation.¹⁶ The ground of the I-thought, my first-person experience of ‘wanting X’, must be ineliminable by the third-person perspective of what *causes* my ‘want’. It is ‘I’ that wants X and not a drive or drive-relation. The self-conscious unity of the ‘I’ that is inherent to practical agency does not permit a self-conception in terms of drives, even though our actions can be causally explained via relations between drives. This draws a clear boundary between the third-person composites and the first-person ‘I’, and so between the practical and the theoretical, which arguably resolves the ‘lack of fit’ and preserves the first-person perspective required for practical agency.

2. The Self in Nietzsche’s Drive Psychology

Chris Janaway¹⁷, too, argues that there is a ‘lack of fit’ between Nietzsche’s practical and theoretical remarks, and that a strictly naturalist¹⁸ conception of agency falls short of self-knowledge and self-valuation. Nietzsche then cannot eliminate the inherent unity of the self-conscious ‘I’. Contra Gardner, he argues that the unity Nietzsche has in mind for his practical

16 Maudemarie Clark / David Dudrick, Nietzsche on the Will: An Analysis of BGE 19, in: Ken Gemes / Simon May (Eds.), Nietzsche on Freedom and Autonomy, New York 2009, pp. 247-268, approach the will and willpower in this manner, but without the commitments associated here with the transcendental reading. Rather Nietzsche’s metaphor of political and so normative relations between drives allows him room to describe first-person practical agency in third-person terms. We are indebted to their discussion on the important distinction between command and obedience on the one hand, and strength and weakness on the other, as relations between drives, which helped us towards recognising the bridge. Cf. Clark / Dudrick, Nietzsche on the Will, 2009, p. 266.

¹⁷ Chris Janaway, Autonomy, Affect, and the Self in Nietzsche’s Project of Genealogy, in: Ken Gemes / Simon May (Eds.), Nietzsche on Freedom and Autonomy, New York 2009, pp. 51-68.

¹⁸ The accounts set up in opposition here are presumably those that read Nietzsche as an ‘epiphenomenalist’ about the conscious self, e.g. Brian Leiter, Nietzsche’s Theory of the Will, in: Ken Gemes / Simon May (Eds.) Nietzsche on Freedom and Autonomy, New York 2009, pp. 107-126. See also Deleuze’s reading: “In Nietzsche consciousness is always the consciousness of an inferior in relation to a superior to which he is subordinated or into which he is “incorporated”. Consciousness is never self-consciousness, but the consciousness of an ego in relation to a self which is not itself conscious.” (Deleuze, Nietzsche and Philosophy, Hugh Tomlinson (Trans.), London 1983, p. 39. A similar reading and emphasis on the self as the unconscious operations of the body can be found in Klossowski who claims: “The *Selbst*, for Nietzsche, has a double meaning: on the one hand, it is, morally speaking, the *Selbstsucht* (the greediness of the self, which is erroneously translated as ‘egoism’), and on the other hand, it is force, unconscious to the cerebral consciousness, which obeys a hidden reason.” (Klossowski, Nietzsche and the Vicious Circle, Daniel W. Smith (Trans.), London 1969, p. 32.

exemplar is a distinct kind of unity not achievable by everyone: it is “a task reserved for a few of us rather than a given”.¹⁹

The main inconsistency, for Janaway, can be found in the relation between Nietzsche’s description of the individual, i.e. what the individual is, and what he prescribes for the individual, i.e. what is exemplary of agency. Nietzsche construes the individual as a complex of will-like components, i.e. drives, which form certain relations with one another. Hence,

Given that this plurality of will-like striving components held in tension is what I am, it makes sense *prima facie* to hold that if I am to *know* anything, it must be through some activity of the drives that compose me and the feelings essentially involved in their activity.²⁰

A drive is “a relatively stable tendency to active behaviour of some kind” and an affect is “what it feels like when a drive is active within oneself”.²¹ An affect is thus a consequence of a drive’s activity or interaction between drives.²² It is categorized in terms of inclination and aversion—a disposition towards or away from something. Affects are experienced as always in relation to some ‘*thing*’ in the broad sense, whether a concept, a topic or an object.²³ Without the relation between thing, affect and drive, and without the implied multiplicity of affects and drives in relation to a thing, there is no knowledge or objectivity. To know something, including oneself, is to do so through drives and affects, which colour our experiences. The basic unit or composite of the self or agency in this reading then is the ‘drive’.²⁴

¹⁹ Janaway, *Autonomy, Affect and the Self*, p. 65.

²⁰ Janaway, *Autonomy, Affect and the Self*, p. 56

²¹ Janaway, *Autonomy, Affect and the Self*, pp. 55-6

²² Janaway, *Autonomy, Affect and the Self*, p. 52

²³ Cf. Janaway, *Autonomy, Affect and the Self*, p. 54

²⁴ We have followed Janaway in avoiding the discussion of construing the drives as “miniature subjects” or “under-souls” Janaway, *Autonomy, Affect and the Self*, pp. 55-6. This discussion is not immediately relevant for our purposes, but a note merits its place here to explain why we postulate what we are calling ‘the bridge’. We cannot correctly attribute a self or the unity of the self to any *individual* drive. We can construe the self as an order, formation, or interaction between drives, but each drive cannot be construed as a self, for we are then moving the complexity of the self as a multiplicity of drives and their interactions one drive. To have a self, in the only sense we know (in third-person terms), is to have a multiplicity of drives. But, this is nothing special or new. We employ the same reasoning when we consider the compositions of other objects. Let us take the classic example of the chair to illustrate this point. In certain contexts, primarily the scientific, we construe the chair as a multiplicity of atoms or subatomic motions with certain relations, **but we do not construe each atom as a chair-**

The above characterises Nietzsche's theoretical account of agency. His practical account depicts a different picture. The agent is construed as active and self-conscious, and who must possess and have power over his or her drives and affects:

To have one's pro and contra in one's power is to make one's knowledge more 'objective'; the plurality of affects, the greatest possible difference in affective interpretations, is 'useful' for knowledge and makes it more 'complete'.²⁵

Relevant for agency in the above quote is the notion of 'having one's pro and contra in one's power', and what is implied by 'having in one's power'. In *BGE* 284, Nietzsche speaks of "freely having", "to *seat* oneself on them as on horses", and "know how to employ". Similar claims are present in *HH I* Pref. 6, where he speaks of "becoming master over yourself", and asserts that "you shall get control over your for and against". These passages purportedly lead Nietzsche to dissolving "the self into a multiplicity of affects and drives", and sticking to "his

atom. Other than being objects composed of atoms, chairs have certain practical functions or aesthetic properties. Thus we cannot describe or construe the part in terms of the whole, i.e. the atoms that compose a chair in terms of chair-atoms, or the drives that compose the self in terms of self-drives. We can construe the whole and how it works through the relations between its parts, but we cannot reduce nor eliminate the whole in favour of its parts. We can certainly imagine calling an atom a chair-atom in certain contexts where the boundaries between different objects are highlighted and analysed: we can say that this atom here is a chair-atom and not a table-atom, i.e. it belongs to the chair and not to the table. In order for the previous to make sense a relation of identity, or an act of identification, must take place between the particular atom and the chair, and so between the one perspective on an object and another; e.g. this atom here X is identified with and belongs to this chair Y and not to another object Z, i.e. the table. This act of identification we are calling the 'bridge', which shows us that there is only a difference in perspective between the self as that which utters "I this or that" and the complex of drives. Our language shows us only a difference in perspectives: how we transition from one to the next and how we employ both towards different ends. Nothing more than that. What then is a chair? It is something we ordinarily construct, trade, sit on etc., but also a complex of atoms, both of which are two different perspectives on the same thing. The question we ought to ask is how does, or how can, one possibly reduce or eliminate a perspective? The same reasoning occurs when we think of and talk about persons. George is a gentleman who adores the Oriental cuisine and Beethoven, but some of us are also aware that he is also a body with cells and organs that operate on electro-chemical reactions, which lead to relatively stable and recurring behaviours construed as drives. When we finally converse with him, however, we notice that the latter picture sinks from our focus and we approach him as a person with desires, hopes and intentions. Nietzsche cannot eliminate this difference in perspective, nor, we argue, does he aim to do so.

²⁵ Janaway, *Autonomy, Affect and the Self*, p. 52.

aims at improving our capacity for knowing and skilfully using our affects”.²⁶ If the ‘I’ and ‘self-knowledge’ are *only* relations between drives and affects, then how can I control or possess them? Possession implies a distinction and separation between the possessed and the possessor, and a unity in the possessor rejected by Nietzsche.²⁷ I cannot confuse myself with my possessions if I am to call them ‘mine’.²⁸ If affects A and B both speak interchangeably about a subject matter and I am to harness, possess or identify with either or both, then I require a self-conscious unity that is distinct from what is unified. Consequently, for Janaway, the previous is only possible for “unified self-conscious subjects, subjects of ‘I’-thoughts”.²⁹

This need not introduce the metaphysical implications that Nietzsche is determined to avoid; the unity of the self-conscious ‘I’ is not a *metaphysical substratum*. Nevertheless, the unity of the self-conscious ‘I’ is indispensable to preserving the reality of first-person practical agency. The main difference between Gardner’s and Janaway’s reading is that the former construes self-conscious unity as inherent to agency *as such*, whereas the latter sees

²⁶ Janaway, *Autonomy, Affect and the Self*, p. 60.

²⁷ See also Chris Janaway, *Morality, Drives, and Human Greatness*, in: Chris Janaway / Simon Robertson (Eds.), *Nietzsche, Naturalism and Normativity*, Oxford 2012, pp. 183-201; see especially pp. 192-195 for a further discussion of why Nietzsche requires the self-conscious ‘I’.

²⁸ The point highlighted by Janaway is regarding thoughts of self-identification and self-possession, which he argues cannot be reduced to relations between drives. We believe the problem is better posed by asking whether in construing people in terms of drive-relations, we risk losing first-person practical agency? We aim to show that Nietzsche’s proposed conception of genuine agency avoids this loss. We can construe the ‘I’ and the content it identifies with (the ‘I’-thoughts) as the majority in relation to the minority using Nietzsche’s political metaphor (BGE 19). Interactions between I-thoughts can be also construed as interactions between an already established order of drives and an activity of drives that disrupt that order. We ought to be clear, however, that this merely shifts Nietzsche’s understanding of the ‘I’ from one perspective to another. It does not in any way commit him to a loss of the ‘I’-perspective, or its unity. We are speaking about the same thing, person and activity here and not two different things, people or activities. Simply put when you ‘look’ into me and describe what you see, it is still *me* you are looking into and not something or someone else. Though the ‘I’ is inherently unified by virtue of what he calls the seductions of language and grammar, its content, i.e. its affective states and action-tendencies (or even thoughts), are nevertheless not necessarily so. The ‘I’’s inherent unity then is merely that of a practice or a game, that sometimes accords with its content and sometimes not, but in all cases it refers to that content and to nothing else—certainly not to some special, simple and unrelated unity or *thing in itself*. For a more elaborate and alternative discussion of Nietzsche’s third-person conception of the ‘self’ see Lanier Anderson, *What is a Nietzschean Self?*, in: Chris Janaway / Simon Robertson (Eds.), *Nietzsche, Naturalism and Normativity*, Oxford 2012, pp. 202-235.

²⁹ Janaway, *Autonomy, Affect and the Self*, p. 60.

the sovereign individual's agential unity as an achievement and task set for the few.³⁰ Janaway shows us a distinction between two types of unity: one inherent to practical agency *as such* that requires the self-conscious unity of the 'I', and the other of the sovereign individual that represents a unity constitutive of self-knowledge. We agree with Janaway that practical agency requires the self-conscious 'I', but agency *as such* for Nietzsche does not require self-knowledge or even consciousness.³¹ Janaway is also right to claim that self-knowledge distinguishes the sovereign individual from ordinary agents, but this is insufficient to account for the sovereign individual's distinct agential unity. The sovereign individual does not only have self-knowledge, but also self-mastery³², which is arguably more than having knowledge of one's drives and affects.³³

3. The Challenges of the Transcendental Reading

Both Gardner and Janaway's readings of Nietzschean agency are right to argue that we cannot eliminate the 'I' without also eliminating first-person practical agency. They are also right to argue that there is a lack of fit in our best interpretations of Nietzsche, if not his own thoughts.

30 See for example: "none of this shows that self-knowledge is impossible: only that it is rare among human beings, that it is a task set for a few of us rather than a given, and that its achievement is a matter of degree" (Janaway, *Autonomy, Affect and the Self*, p. 65).

31 Cf. GS 354: "For we could think, feel, will, remember, and also 'act' in every sense of the term, and yet none of all this would have to 'enter our consciousness' (as one says figuratively)".

32 See for example Nietzsche's description and the language he uses for the sovereign individual: "... and how could he, with his self-mastery, not realize that he has been given mastery over circumstances, over nature and over all her creatures... The proud knowledge of the extraordinary privilege of responsibility, the consciousness of this rare freedom and power over himself and his destiny, has penetrated him to his lowest depths and become an instinct, his dominant instinct..." (GM II 2). We quote Friedrich Nietzsche, *On the Genealogy of Morality* (GM), Keith Ansell-Pearson (Ed.), Carol Diethe (Trans.), New York 1997, p. 37. The description here is saturated with words that indicate not only self-knowledge but also self-mastery; self-knowledge is thus necessary but insufficient for self-mastery; see GS 17, 78, 79, 99, 290 & 299 for more on the relationship between self-knowledge and self-mastery.

33 Robert Pippin, *How to Overcome Oneself: Nietzsche on Freedom*, in: Ken Gemes / Simon May (Eds.), *Nietzsche on Freedom and Autonomy*, New York 2009, pp. 69-88, also contends that the sovereign individual is "a complete and hierarchical unity among states of one's soul, memories, desires, aversions, and so forth" (p. 77). He adds that it "is a distinct sort of psychological self-relation, both attitudinal and dispositional—in order to be capable of any real practical intentionality or real agency" (p. 79). He further adds that this unity requires us to assume what Nietzsche is attempting to avoid, namely an "independence between a creator or ruler self and a created object and 'commonwealth'" (p. 78), which unfortunately leads us back, full-circle, to the lack of fit.

They are wrong, however, in claiming that rejection of the causal role of the 'I' entails its elimination.³⁴

The transcendental reading reveals two difficulties in Nietzsche's claims on selfhood. First, the practical and theoretical conflict due to the presumed theoretical elimination and simultaneously the practical commitment to the 'I'. Second, the claim that the unity of the self-conscious 'I' is *wholly* distinct from the drives but *somehow* arises from **the drives or some drive-relation** **? anything what ?** requires explanation. The account can thus be summarised as follows:

1. The third-person perspective on agency shows a complex multiplicity of drives brought into unity by hierarchical power relations which generate actions.
2. The first-person perspective shows a simple unity and sense of ownership and control over our actions.
3. The simple unity of first-person practical agency cannot be explained via drives; it grounds practical agency, and this is attributed to the self-conscious 'I'.
4. By being inherent to first-person practical agency, the self-conscious 'I' is distinct from the third-person drives and ineliminable.

The concern here is with the move from 1 to 2. The transcendental reading contends that 2 cannot be explained by a naturalist and strict third-person account of selfhood. Relations between drives cannot account for the unity of first-person practical agency. They contend that 3 and 4 must follow to explain the relation between 1 and 2.³⁵

The transcendental reading is correct in claiming that first-person practical agency requires the 'I', and that insofar as we engage in first-person practical agency we cannot eliminate it. One cannot *fully* construe nor relate to oneself as a set of drives.³⁶ An agential

34 See especially Gardner, Nietzsche, the Self and Disunity, pp. 6-7, and Chris Janaway, Beyond Selflessness, New York 2007, p. 217.

35 For an alternative discussion and solution to the lack of fit than we have presented here, but which accords with our objections to the transcendental reading, see Anderson, What is a Nietzschean Self?, pp. 203-208.

36 There are some thinkers, i.e. Paul M. Churchland, Eliminative Materialism and the Propositional Attitude, in: The Journal of Philosophy 78 (1981), pp. 67-90, who hold an eliminativist materialist position by arguing for the redundancy of first-person self-expression and self-reference, which requires revisions in how we communicate about ourselves and our experiences. Churchland's eliminativism entail's a revision of language, **i.e. adding and removing linguistic expressions (pp. 88-9)**, in favour of a more accurate explanation and prediction of behaviour. Materialistic descriptions then should take precedence because they purportedly better explain behaviour. The

action can be causally-explained in terms of drives and affects in the third-person, e.g. John bought flowers and asked Julie on a date because his sexual drives (coupled with other drives) were active in her presence. John, however, is also aware that *he* bought flowers and asked Julie out, *not* a set of drives. John's attribution of his affects and actions to himself as opposed to a drive or drive-relation, cannot be explained by the drives. Nietzsche must therefore accept the 'I', along with the belief in it, for the sake of first-person practical agency, in which his sovereign individual must also engage.³⁷

It is vital to note, however, that Nietzsche's critique of the 'I' is not concerned with its utility, but with whether or not it is *sufficient* to be the genuine determinant and source of our actions and values. In other words, Nietzsche poses the following question: is simply saying 'I will do this or that' sufficient for the deed to ensue? Can its inherent unity be attributed to that

aim of the eliminative materialist is therefore accuracy in the prediction and explanation of behaviour. It is dubious that the same aim can be attributed to Nietzsche (cf. GS 121, BGE 24, Nachlass 1885, 34[253], KSA 11.506; Nachlass 1885, 38[3], KSA 11.597f.; Nachlass 1886-1887, 7[63], KSA 12.317). His aims are self-affirmation and the enhancement of life with focus on how our ideas about life, morality, reality and ourselves affects the affirmation and enhancement of life. Whether these aims can be achieved by an eliminativist materialist proposal about the redundancy of linguistic practices is something that Nietzsche would have to deny. He criticizes the will to truth because the conditions of life may *require* error or lies (cf. HH I Pref. 1, GS 121 and BGE 16).

37 Chris Janaway, Nietzsche, the Self, and Schopenhauer. in: Keith Ansell-Pearson (Ed.), Nietzsche and Modern German Thought, London 1990, pp. 119-142, is right to claim that we cannot speak of agency without in some sense presupposing the subject: "we must still use the concept subject even if not of any single entity" (p.135). Nevertheless, given the possibility of errors in self-knowledge, using the concept is not sufficient to ground the assumptions that the concept makes about agency, i.e. its presupposing the sort of ownership and control "that gives style, imposes accord, identifies with all of its actions, harnesses or controls the many drives and affects" (p. 137). That they are mine or attributed to me is some paces away from my being able to control them; control, or self-mastery as Nietzsche calls it, is not given by the 'I'. Janaway is right to argue that the 'I' is inherent to first-person practical agency, but we ought to be careful not to give the concept 'I' a more privileged status than that of a concept, i.e. a way of representing the relationship between our resolve and our actions, which may or may not be as unified as the inherent unity of the 'I' suggests. We must also be clear that agential unity rests on a lot more than the use of the concept 'I'. The 'I' is not sufficient for agency *as such*, though it is necessary for first-person practical agency. The unity of the 'I' and the unity of agency (the relationship between one's actions and one's will or resolve) ought to therefore be distinguished. In other words, we have to take seriously Nietzsche's claim that the concept of the subject or the 'I' is trapped in the seductions of language and grammar—more work and discussion can go on the relationship between the 'I' and language, in Nietzsche. For a collection of essays that seek to address this see João Constâncio / Maria João Mayer Branco (Ed.), *As the Spider Spins: Essays on Nietzsche's Critique and Use of Language*, Berlin / Boston 2012.

to which it refers, i.e. agency and so the relationship between one's resolve and one's actions. Nietzsche is probing us to reconsider whether our actions are as unified as our 'I'-claims suggest and whether 'we', the first-person 'I', are the true causal determinants of our actions. This is easier to follow if we consider his arguments against self-knowledge. Since we can err in what we know about ourselves and thereby act and communicate about our actions in accordance with this error, then we cannot raise first-person practical agency to the seat of agency *as such*. The 'I' is therefore insufficient for a complete account of agency, which should include or explain the errors in self-knowledge, or the lack of self-knowledge.³⁸

If the inherent unity of the 'I' would be sufficient for agency, then, irrespective of the relationship between our resolve and our actions, we would all practice genuine agency. What we think about ourselves and our values, however, is not always correct or even correlated by our actions. The relationship between first-person practical agency and weakness of will leads Nietzsche to offer an alternative account of genuine agency that rests on the relationship between one's resolve and one's actions. Since the 'I' refers equally to both a weak and a strong will, its inherent unity can be misleading by suggesting the kind of unity constitutive of a strong will, i.e. a will whose resolve and actions are in harmony. It is on this basis that he construes the 'I' as a fiction or illusion. This shows us then that Nietzsche is not questioning the 'I' *as such*, but its authority independent from its correlating actions. He challenges us to act in accordance with what we think and reason about ourselves and to give our actions the kind of unity we profess to have in first-person practical agency. Notice, however, that this does not commit him to reduction or elimination of first-person practical agency, only to a critique of the scope of its inherent unity and its legitimacy with respect to genuine agency.

4. Towards an Alternative Solution

38 Paul Katsafanas, Nietzsche on Agency and Self-Ignorance, in: *Journal of Nietzsche Studies* 43 (2012), pp. 5-17, offers a very useful and detailed discussion of the role of self-knowledge in Nietzschean agency. We agree with him that Nietzsche's views on agency mounts an attack on previous conceptions, with a view towards a new and unique account of what constitutes genuine agency. However, Katsafanas's interpretation still leads us to the lack of fit. This is because although he rightly rejects epiphenomenalism and its opposite, he still holds the view that we argue Nietzsche is trying to avoid: a kind of interactive dualism between drives and conscious thought, deemed to be, and in fact must be, an interaction between two distinct things, which leans on conscious thought by contending that genuine agency can only be understood in terms of the central role that conscious thought plays in our actions.

There are two implicit commitments in the transcendental reading's solution to the lack of fit that are not sympathetic to Nietzsche's critique of agency, and leads us to misconstruing his aims of offering an alternative account of genuine agency.

The first commitment can be located in Gardner's proposal, which claims that the unity of the self is *grounded* in the self-conscious 'I'.³⁹ We will call this the 'false identity commitment'. The use of ground here is unclear. If by ground we mean the ground of first-person practical agency, then Gardner's challenge is correct. If by ground, however, we mean the ground of agency *as such*, then Nietzsche's critique is illuminating. We do not know ourselves as well as we claim, and we moreover do not have the kind of mastery over our actions suggested to us in first-person practical agency (see D 105 & 115; also BGE 17 & 19). We can err in what we assume to be driving our actions (see HH Pref. 1). For example, Julie may think of herself as a very altruistic and generous person, who will help a friend in need, and she does. Later she experiences a situation where she is now in need, but the same friend does not help her. She responds with anger and scornfully reminds him of the time she helped him. Later she reflects on the situation and is bewildered by her response. She wonders if she is in fact an egoistic person, because only egoism would motive her subsequent expectation of reciprocity and anger at its lack. Nietzsche offers us similar examples of the errors and deceptions of first-person experience (see GM, Pref. 1, and HH, Pref. 3; see also TI, The Four Great Errors), and highlights the limitations of our self-conscious access to the totality of our activities, their motivations and their geneses (see especially D 115).

Nietzsche is, we believe, making the following claim: if we can err about our motivations and our self-image, then we cannot assume that the unity of the 'I' is the ground of agency *as such*, even though the 'I' itself is indispensable to first-person practical agency.

The false identity commitment incorrectly establishes a relationship of identity between the unity of the 'I' and the unity of agency *as such*, which Nietzsche is rejecting in the majority of us.⁴⁰ **We assume that we are unified and that we act in accordance with this**

39 See also Anderson, What is a Nietzschean Self?, p. 213, for a similar objection to Gardner's reading.

40 Paul Katsafanas, The Concept of Unified Agency in Nietzsche, Plato and Schiller, in: Journal of the History of Philosophy 49 (2011), pp. 87-113, argues that conscious thought and reflection are necessary for distinguishing between ordinary and genuine agency. He contends that Nietzschean agential unity is best characterised by the relationship "between drives and other parts of the individual", namely, "self-conscious thought" (p. 103). The lack of fit problem shows us, however, that Nietzsche cannot explain how self-conscious thought can work given our third-person identification with drives. The claim that the agent is not only a set of drives, but also has other parts, bypasses the problem of the lack of fit, but does not resolve it. It moves its

unity, but this is undermined by the many cases in which we act in discordance with our self-image. Nietzsche's unique account of genuine agency then begins without the premise that the unity of the self-conscious 'I' is synonymous with agency *as such*, or that the unity of the 'I' is somehow imposed on our actions. Therefore, first-person practical agency is necessary, but insufficient for genuine agency. Nietzsche's practical exemplar requires *more* than the ability to speak and think in terms of an 'I', or participate in first-person practical agency, to be practising *genuine* agency.

The second commitment follows from the previous by holding that the unity of the self-conscious 'I' is distinct from and cannot be explained by the drives; it is irreducible and ineliminable.⁴¹ The unity of the 'I' is inherent to agency, but separated from any propositions about it in third-person terms. We will call this the subjectivity commitment. This commitment is often used to resolve the lack of fit, but leads to shifting its burden. This shift of burden can be best captured by the following question: how can something be constituted by drives but not be explicable in their terms, or how can it possibly *interact* with the drives without, in some way, being the same kind of *thing*, i.e. a drive or drive-relation?

burden elsewhere. We now require a story that explains the interaction between drives and self-conscious thought, just as we require one for the 'I' as a faculty and the drives. It is unfortunate that Katsafanas did not explore a solution to this, cf. Katsafanas, *Unified Agency*, p. 113, footnote 57.

41 Paul Katsafanas, *Nietzsche's Theory of Mind: Consciousness and Conceptualization*, in: *European Journal of Philosophy* 13 (2005), pp. 1–31, pp. 11-13 is also aware of this commitment as well as of the central issue of the relationship between consciousness and the drives. He contends that the problem of consciousness arises from representing it as a faculty of awareness. What Nietzsche really denies is not consciousness per se, but the approach to consciousness as a faculty (p. 12). We agree with Katsafanas that Nietzsche does not seek to undermine consciousness as such, but its being conceived as faculties. This resolves the problem of how conscious agency is possible, and gets us as far as saving first-person practical agency. This does not resolve the lack of fit, however. If self-conscious thought is not a drive nor a faculty, and not even a thing or object, then how can it interact with the drives without also being "in the trenches with the drives" (John Richardson, *Nietzsche's Freedoms*, in: Ken Gemes / Simon May (Ed.), *Nietzsche on Freedom and Autonomy*, New York 2009, pp. 125-149, p. 137)? Consciousness, in this reading, remains a phantom that somehow effects the basic unit explanatory of relatively stable behaviour. The phantom arises, we suggest, because we do not notice that Nietzsche is thinking in terms of a bridge between first and third-person perspectives on agency. He does not construe them as things distinct in kind that somehow interact with and affect one another, but as two conceptions of the same thing or activity, one of which can be misleading with respect to that activity. For an alternative and very useful analysis of Nietzsche's conception of consciousness, conceptual content and the 'self' see João Constâncio, *Consciousness, Communication, and Self-Expression: Towards an Interpretation of Aphorism 354 of Nietzsche's The Gay Science*, in: Constâncio / Mayer Branco (Eds.), *As the Spider Spins*, pp. 197-231.

The subjectivity commitment leaves Nietzsche with the following undesirable impasse. All agential actions are constituted by or are the consequences of something that is indiscernible in the third-person, whilst simultaneously it is accessible to us through the first-person as the unity of the ‘I’. There is *something* in one perspective that cannot be located nor perceived in the other. Nietzsche’s critique of agency can, as we have shown, accommodate the ‘I’ as a practice, but must abandon its inherent unity if it attributes it not *just* to how we think and speak about ourselves, but to agency *as such*, i.e. to the relationship between one’s resolve and one’s actions. Nietzsche’s rejection of metaphysical dualism and his insistence on the identity of the will (self, or agent) with the body inherited from Schopenhauer⁴², entails that he must both defend and illustrate the following: the ‘I’ and what occurs in first-person practical agency is in some way accountable by a manifold of third-person drives. Such an account, we argue, is available to him, but only on the basis of what we call the bridge (more on this below).

The key requirement for the transcendentalists is that Nietzsche shows us how the above is possible without elimination of the first-person perspective, which they argue is unavailable to Nietzsche’s theoretical account. But, as we have seen, it is unclear whether elimination here is of the use of the ‘I’ which *refers* to the relationship between one’s resolve and one’s actions, or elimination of the *unity* between one’s resolve and actions that he denies in the majority of us first-person users of the ‘I’. This lack of clarity, we suggest, constitutes the central limitation of the transcendental reading.⁴³ If the individual in the third-person is a

42 See for example Z I, The Despisers of the Body: “body am I through and through, and nothing besides; and soul is just a word for something on the body.” We quote Friedrich Nietzsche, *Thus Spoke Zarathustra: A Book for All and None (Z)*, Adrian Del Caro and Robert B. Pippin (Eds.), Adrian Del Caro (Trans.), New York 2006, p. 23. For an account of Schopenhauer’s will-body identity see John E. Atwell, *Schopenhauer: The Human Character*, Philadelphia 1990, pp. 27-42. For a thorough comparison between Schopenhauer and Nietzsche on drives and consciousness see João Constâncio, *On Consciousness: Nietzsche’s Departure from Schopenhauer*, in: *Nietzsche-Studien* 40 (2011), pp. 1-42.

⁴³ An alternative reading of the Nietzschean self that can be set up in opposition to the transcendental reading, is the Humean bundle theory of the self. Hume grounded his conception of the self on the claim that the “self or person is not any one impression, but that to which our several impressions and ideas are suppos’d to have a reference” (David Hume, *A Treatise of Human Nature*, David Fate Norton / Mary J. Norton (Eds.), New York 2000, p. 164). He then proceeded to argue that no such idea or impression exists, but we have several such ideas and impressions. Consequently the self or person is “nothing” but a bundle or collection of different perceptions, which succeed each other with an inconceivable rapidity, and are in a perpetual flux and movement. Our eyes cannot turn in their sockets without varying our perceptions. Our thought is still more variable than our sight; and all our other senses and faculties contribute to this change; nor is there any single power of the soul, which

multiplicity of drives, and in the first-person she practises a unity that owns and is distinct from her drives, then there must be something parallel to this unity in the drives—a state of affairs, relation, or activity—that corresponds to the first-person unity, even if it is not reducible to it. The claim should be construed as offering two different perspectives on the same ‘thing’ or activity, which would make his claims on genuine agency coherent without entailing a third-person *phantom*. Nietzsche indeed cannot claim that one perspective has a truth-value which negates or renders the other superfluous, but must show us a bridge between the two that makes his third-person analysis and conception of the ‘self’ coherent. Elimination is thereby justifiably out of the question.

We agree with the transcendental reading that Nietzsche requires an account of the self-conscious ‘I’ that fits with first-person practical agency. This account, however, must tell us more than that the ‘I’ is a unity separate from or unrelated to our other experiences, and that we *are* this unity. We seemingly have two options. Firstly, the ‘I’ is some ‘thing’, but the kind of thing that we cannot perceive, which entails a commitment to metaphysical dualism and reduces the ‘I’ to a third-person phantom.⁴⁴ Second, the ‘I’ is not a ‘thing’, but a practice we engage in. The ‘I’ becomes apparent as a problem only when we ask ‘what is it?’, and

remains unalterably the same, perhaps for one moment. The mind is a kind of theatre, where several perceptions successively make their appearance; pass, re-pass, glide away, and mingle in an infinite variety of postures and situations” (p. 165). A discussion on the relationship between Hume and Nietzsche’s conception of selfhood is very useful for showing how Nietzsche fits in the history of philosophy. A comparison between Humean passions and Nietzschean drives would also be useful to unpacking some of the limitations of a strictly naturalist reading of Nietzschean agency. This is beyond the scope of the current paper, however, and so we will settle for pointing to some papers that directly or indirectly deal with the issue as a first-step towards a broader discussion on the topic. For a Humean reading of Nietzschean selfhood that is also conscious of the limitations of a Humean approach to the Nietzschean self can be found in Anderson, *What is a Nietzschean Self?*, pp. 202-235. Anderson rightly claims that Nietzsche’s conception of selfhood sounds similar to a Humean bundle theory, but emphasises a key difference: “drives and affects are themselves standing attitudes that persist, rather than fleeting, occurrent states *à la* Hume” (p. 224). For an alternative and very compelling argument on why Nietzsche’s conception of the self is distinct from Hume’s see Gardner, who argues that Hume’s theoretical-practical disunity differs from Nietzschean disunity by virtue of their differing philosophical projects, “the enlightened Nietzschean subject, who does not receive values passively from nature but is bound to *innovate* them, and who is consequently exposed to the full battery of reflective questions which, in the Humean subject, the operation of the passions obviates the need and leaves no scope for” (Gardner, *Nietzsche, the Self and Disunity*, p. 17). For a more detailed comparison between Hume and Nietzsche see Brian Leiter, *Routledge Philosophy Guidebook to Nietzsche on Morality*, London 2002.

begin to locate that which corresponds to it. Though we do not and cannot possibly do so, it commits us to believing in it because we use it. Both alternatives leave us with a kind of phantom, which Nietzsche is trying to avoid. But to succeed in avoiding it, he requires a bridge between first-person practical agency and the third-person manifold that preserves his critique of both transcendentalist and naturalist accounts of agency⁴⁵, along with his conception of genuine agency.

The transcendental reading loses sight of Nietzsche's alternative conception of genuine agency, and the fact that this commits him to avoiding reduction or elimination. This is patent in his rejection of naturalist hypotheses that are 'causal mechanistic'⁴⁶. Not all third-person accounts of agency need be 'causal-mechanistic' and so fall short of representing the 'I'. The 'I' is not sufficient for genuine agency, but we must not construe it and that to which it refers as separated from the third-person perspective. Consequently, to preserve the 'this-worldliness' of the agent or self, but also not lose an important practice, Nietzsche requires a 'bridge' between the two perspectives.⁴⁷ If what we are calling the bridge here is implausible,

44 See Richard Schacht, Nietzsche, New York 1983, pp. 131-140, for a detailed discussion of why Nietzsche is adamant to reject this solution.

45 It is important to note here that our rejection of the transcendentalist reading does not entail that we accept a strict causal-naturalist or even epiphenomenalist reading of Nietzschean first-person practical agency, defended by Brian Leiter, The Paradox of Fatalism and Self-creation, in: Brian Leiter / John Richardson (Eds.), Nietzsche, New York 2001, pp. 281-321. See also Brian Leiter, Nietzsche's Theory of the Will, in: Leiter /Richardson (Eds.), Nietzsche, pp. 107-126, for a more recent and in many respects similar position. We argue conversely that Nietzsche is aiming for and requires a middle-ground between the drives and 'I'-thoughts, which must accommodate the theoretical value of a drive and the practical value of the 'I'. What we have called the bridge between the first-person and third-person here is an attempt at such middle-ground.

46 Cf. BGE 12: "Between you and me, there is absolutely no need to give up "the soul" itself, and relinquish one of the oldest and most venerable hypotheses – as often happens with naturalists: given their clumsiness, they barely need to touch "the soul" to lose it".

⁴⁷ For a very useful and informative attempt to resolve the lack of fit see Constâncio, On Consciousness, pp. 1-42. Constâncio aims to resolve the lack of fit by ascribing 'perspectives' to drives in accordance with what the "continuum model" (pp. 21-26). He claims that "as perspectival valuations, the drives are always changing and adapting to their surroundings — so that at every moment the "total state" of the organism is a cluster of perspectival relations, not an aggregate of mechanical, causal relations among atomic parts" (p. 23). We agree with him that causal interaction between the 'I' and the drives is what Nietzsche is trying to undermine for precisely the reason that it commits him to some kind of atomism (and brings back what we are calling the phantom), which he seeks to avoid. We disagree that Nietzsche could replace causal interactions with perspectival interactions between the drives, or that Nietzsche sought to ascribe a perspective to a drive in any strict sense. Our disagreement rests on the homunculi problem, i.e. the drives as miniature agents, which is

then he must adhere to some form of dualism that affords privileged status to first-person practical agency. This would require him to accept that something participates in, influences or even determines our actions that we cannot perceive nor conceive from any other perspective than the first-person, wherein the most we can conceive about it is that it shows a distinct kind of unity, ownership and identification with our actions. (It is important to highlight here that, for our purposes, Nietzsche does not have to give us a positive account of what the third-person parallel to the 'I' is, but rather that he commits himself to it and thinks on the basis of this commitment; alternatively, his thoughts on genuine agency would rest wholly on a phantom, because, to make practical sense, i.e. to show how his genuine agent would relate to herself and speak to the rest of us about herself, they required what the phantom represents, i.e. a unity inaccessible to the third-person that has a third-person effect.)

In short, the lack of fit cannot be resolved unless we recognise that Nietzsche's aim is to critique the 'I's supposed relationship to our actions and their geneses (see TI, The Four Great Errors), and to critique its influence on our actions and evaluations (D 109, 115, 116,

usefully discussed in Paul Katsafanas, On Homuncular Drives and the Structure of the Nietzschean Self, in: The Journal of Nietzsche Studies 45 (2014), pp.1-11. Constâncio's suggestion that we should construe the conscious agent or the conscious mental state as the 'total state' of the organism, however, is useful for the lack of fit debate, cf. (WLN: 1[61]): "Every thought, every feeling, every will is *not* born of one particular drive but is a *total state*, a whole surface of the whole consciousness, and results from how the power of *all* the drives that constitute us is fixed at that moment — thus, the power of the drive that dominates just now as well as of the drives obeying or resisting" (Friedrich Nietzsche, Writings from the Late Notebooks (WLN), Rüdiger Bittner (Ed.), Kate Sturge (Trans.), Cambridge 2003, p. 60). We should notice in the above, however, that Nietzsche does not claim that each individual drive has a perspective, but that the state of affairs between our drives (their power relations) constitute a total state that is a perspective on this or that thing or circumstance. Also, Nietzsche does not claim that we should replace mechanical or causal relations between drives for perspectival relations, but rather power relations. This accords with the solution we are calling the bridge and is more apt to avoid the homunculi problem by not ascribing a perspective to a drive. It is not this or that drive that takes a perspective, but this or that *total state* that corresponds to, or is identical to, this or that perspective that agents take on this or that thing (in the broad sense). We should be careful here not to project the inherent unity of the 'I' onto the drives, and a perspective is such a projection. Drives need not be unified *in the same way* as the 'I' in order for us to have a perspective on this or that thing or circumstance. Nietzsche is cautious to not overestimate conscious unity when compared with unity of the organism (cf. GS 11); the unity of the 'I' is for him merely a linguistic unity or the unity of a practice that does not always correspond with that to which it refers. There is no drive-independent perspective than this total state, for Nietzsche. Accordingly, a drive does not take a perspective, but *participates* in the formation of a perspective by pushing or pulling towards this or that action or thing. A perspective is the total state relative to this or that thing or circumstance at that moment, which can and for Nietzsche often does change with time.

119; BGE 12, 17, 19). He focuses on the believing, the valuing and the becoming (the supposed content and modifications of the ‘I’) more than the agent (the ‘I’) who believes, values and becomes, not to eliminate the agent or first-person practical agency, for he cannot deny that we in fact *do* think, communicate and act using the ‘I’. On the contrary, he seeks to critique our presuppositions on the role that the ‘I’ plays in agency; not the practical use of the ‘I’, but the assumption⁴⁸ that we are unified, self-conscious subjects irrespective of the state of affairs of our drives. This is done as a preparation for offering an alternative, genuine account of agency as a unity that is willed and achieved, not one that is given. Nietzsche questions our assumption that our actions are modifications of a unified cognitive subject (see BGE 16) that imposes its unity on a manifold of conative states. He repudiates accounts of agency that takes the unified subject as simple, or an atom, or as *given*.⁴⁹

In the following passage from BGE Nietzsche expresses his critique of the inherent unity of the ‘I’ attributed to first-person agency, only here it is approached from an analysis of what occurs when we will or resolve to undertake a course of action.⁵⁰ This approach, we

⁴⁸ Nietzsche often construes this assumption as arising from grammar and language. Cf. TI, ‘Reason in Philosophy’, 5: “Language began at a time when psychology was in its most rudimentary form: we enter into a crudely fetishistic mindset when we call into consciousness [*Bewußtsein*] the basic presuppositions of the metaphysics of language — in the vernacular: the presuppositions of reason. It sees doers and deeds all over: it believes that will has causal efficacy: it believes in the ‘I’, in the I as being, in the I as substance, and it projects this belief in the I-substance onto all things — this is how it creates the concept of ‘thing’ in the first place ...”.

⁴⁹ For a useful and elaborating discussion on Nietzsche’s critique of agential unity see Alexander Nehamas, *Nietzsche: Life as Literature*, Cambridge, Massachusetts 1985, pp. 176-190. See also Robert Pippin, *Nietzsche, Psychology, and First Philosophy*, Chicago 2010, pp. 75-82, for an alternative discussion.

⁵⁰ Paul Katsafanas, *Nietzsche and Kant on the Will: Two Models of Reflective Agency*, in: *Philosophy and Phenomenological Research* doi: 10.1111/j.1933-1592.2012.00623.x, is right to show that Nietzsche transitions between various uses of the word will. Often and in the earlier period of his writing, he construes the will’s freedom relative to whether or not it is causally determined. In the latter period, he alters his criterion for construing the will in terms of whether or not it can overcome some resistance (internal or external) that arises with respect to the willed activity, i.e. whether it is strong or weak. The latter approach to the will shows us that Nietzsche has in mind an alternative account of agency. Nietzsche requires a story, however, which explains why and how “passion and reason are both efficacious” (p. 33). This is because if we run with the separation of reason from the passions by virtue of their nature or activity, we end up with causal interaction between two distinct things that somehow do not meet on a common ground, which brings about the kind of phantom we tried to illustrate. This reveals a core philosophical problem. We believe that Nietzsche can escape it only if he accepts the bridge between the first-person and third-person perspectives, which does not rely on causal interaction, but rather on a correlation or identity. In addition, if he argues for a different understanding of the self, which is no longer misleadingly construed as a subject (premised on grammar and language), but as an

believe, is crucial to his position on the *apparent* duality between the first-person and third-person perspectives.

On the one hand, we are, under the circumstances, both the one who commands *and* the one who obeys, and as the obedient one we are familiar with the feelings of compulsion, force, pressure, resistance, and motion that generally start right after the act of willing. On the other hand, however, we are in the habit of ignoring and deceiving ourselves about this duality (*Zweiheit*) by means of the synthetic concept of the “I”. As a result, a whole chain of erroneous conclusions, and, consequently, false evaluations have become attached to the will, – to such an extent that the one who wills believes, in good faith, that willing *suffices* for action. (BGE 19)⁵¹

The interpretation of this passage we are advancing here is unusual, but we believe it to be the best for a genuine solution to the lack of fit.⁵² Nietzsche is patently aware of the apparent discord between the two perspectives, and that it rests on our use of the ‘I’. Notice, however, that he is not objecting to the use of the ‘I’, nor seeking to eliminate it. He is describing how the ‘I’ can affect our idea of ourselves and so self-knowledge at the most fundamental level, i.e. the relationship between willing and acting, and not between ourselves and our drives. It conceals the bridge between the two perspectives, which becomes apparent to us upon analysis of what ensues following our willing a course of action.⁵³ He contends that what

activity, then we have resolved the problem of interaction—namely, there is no interaction. The ‘I’ and the drives are not two different things, but two different perspectives (alternatively two different descriptions) on the same thing, or better yet, activity.

⁵¹ See also GS 127: “With his assumption that only that which wills exists, Schopenhauer enthroned a primordial mythology; he seems never to have attempted an analysis of the will because like everyone else he *believed* in the simplicity and immediacy of all willing — whereas willing is actually such a well-practised mechanism that it almost escapes the observing eye”.

⁵² For an alternative analysis of this passage and its implications for morality see Laurence Lampert, *Nietzsche’s Task: An Interpretation of Beyond Good and Evil*, New Haven 2001, pp. 49-50. For an interpretation that comes very close to the bridge but unfortunately does not avoid the homunculi problem see Clark / Dudrick, *Nietzsche on the Will*, pp. 255-257.

⁵³ Nietzsche reasons on the basis of the possibility and inevitability of a bridge between the two perspectives, e.g. see D 48: “*Know yourself is the whole of science*. – Only when he has attained a final knowledge of all things will man come to know himself. For things are only the boundaries of man.” He does not draw a boundary between the two perspectives, but strongly holds the idea that through the third-person perspective, the agent can attain self-knowledge and self-mastery (see BGE 32). Again, this is not the claim that the ‘true’ self then lies in the third-person and that the first-person is therefore superfluous, but only that it can be useful to the project of self-mastery, i.e. to our willing a course of action and following through. Nietzsche’s contention is that we do

occurs in the process of willing is a distinct kind of activity, i.e. a relation of command and obedience, which is present in both perspectives. In the first-person perspective, it is a command and obedience played within one person. Consider when 'I' deliberate over a course of action and am undecided, or when 'I' struggle to maintain a course of action in light of some alternative temptation, or when 'I' reason and resolve to attend to this and not that, or when 'I' deviate from or even abandon a course of action to undertake a new one. In each case and from the first-person perspective, 'I' am the one who overcomes the temptation, or who alters the course of action, or who feels the resistance when 'I' wish to act contrary to my values or aims, or who is overwhelmed by my own temptation.

In the third-person, however, the same phenomena listed above can be construed as relations of command and obedience between drives and their rank order (*BGE* 6); rank order here refers to the unity (or disunity) of the agent's resolve relative to her actions from the third-person perspective. For example, one's indecision over a course of action represents an unresolved conflict between drives; the deviation from a course of action is the consequence of a disruption of the rank order of drives; the resistance to a particular activity is the direct consequence of a drive(s) becoming active in opposition to the others, and so on. Therefore, closer analysis of the *activity* of willing or resolving to undertake a course of action reveals to us the bridge between the first-person and third-person perspectives that he requires to resolve the lack of fit and make sense of his account of genuine agency. This bridge is the act of will and its corresponding action.⁵⁴

Both perspectives contain the same activity of resolving to act and all that comes with this resolution, but different approaches to it from different perspectives: between one thing and its modifications, or a relation between a multiplicity of things and their order of rank. We describe and construe the *same* thing, event or activity from two distinct perspectives and

not necessarily find or perceive a self, but will a self (cf. HH II, 366). See also GS 333 for more on Nietzsche's commitment to what we call the bridge.

⁵⁴ Rex Welshon, *The Philosophy of Nietzsche*, Chesham Bucks 2004, pp. 152-156, is right to place an emphasis on Nietzsche's conception of the act of will or willing as useful for resolving the lack of fit. We worry, however, that conceiving of an interaction between consciousness and the drives, as he argues, makes it difficult for Nietzsche to have a plausible and coherent solution. We construe the passages where Nietzsche describes the interaction between the drives and consciousness as probative and critical of conceptions of agency that rest on the assumption that genuine agency is determined wholly by the self-conscious 'I'. Thus instead of an interaction between the 'I' and the drives, we suggest that Nietzsche requires an interaction between the 'I' and its actions (in the first-person), or between drive-relations and their corresponding actions (in the third-person). Both refer to the same thing: the act of will or activity of willing. We favour this interpretation of the passage.

in two distinct ways, neither of which can be more right or wrong because both refer to and are grounded in the relationship between the basic *preferred* unit of will or resolve (be it the ‘I’ or a ‘drive’) and its correlate action on the basis of which we both understand and evaluate agency. Problems arise only when we attempt to attribute the unity of the ‘I’ that is inherent to first-person practical agency to its modifications, i.e. its actions **across time**, or, what comes to the same, when we attempt to construe this inherent unity as more than just a practice. **Equal problems will and do arise when we do the reverse and eliminate the unity of the ‘I’ in favour of relations between drives and their rank order.** Naturalist and epiphenomenalist readings who use the ‘drive’ as the basic unit of will or resolve, will encounter similar problems when they try to apply the unity of a multiplicity to that of the ‘I’ in an attempt to preserve the first-person practical agency (though we cannot show this here). Elimination or reduction of the ‘I’ is a consequence of the attempt to dissociate the two perspectives from their bridge, namely the *act of will* and what it entails. This elimination or reduction would be akin to claiming that, upon closer inspection at its composition, the chair in front of us is not a chair because we recognize that it is just arrangement of atoms (or subatomic motions), and that we ought to abandon all talk of chairs and speak only of atoms (or subatomic motions), which is absurd. The fact that *this* chair is *this* composite of atoms is untouched by our calling it a chair and not a composite of subatomic motions in all cases or contexts. The same reasoning operates in our understanding of how the ‘I’ relates to the drives.

Nietzsche analyses and evaluates agency on the basis of how our ‘I’-claims relate to their respective action(s), not solely on the basis of the claims themselves. Genuine agency, **or a normative conception of agency, is based on** the relationship between the preferred basic unit of will and our actions, not *solely* on the basis of our thoughts about our actions.⁵⁵ Since our actions are determined by our drives-relations, genuine agency can in turn *be understood* as the unity of one’s drives relative to one’s actions.⁵⁶ Notice, however, that for us to know whether or not the drives are unified we must look to the coherence and unity of our actions

55 For a discussion on Nietzsche’s prioritization of actions over thoughts about actions see Welshon, *The Philosophy of Nietzsche*, pp. 140-1.

56 For a discussion on the relationship between the unity of drives and genuine agency see, Lanier Anderson, Nietzsche on strength, self-knowledge, and achieving individuality, in: *International Studies in Philosophy* 38 (2006), pp. 89-115; Ken Gemes, Nietzsche on Free Will, Autonomy, and the Sovereign Individual, in: Ken Gemes / Simon May (Eds.), *Nietzsche on Freedom and Autonomy*, New York 2009, pp. 33-50; Simon May, Nihilism and the Free Self, in: Ken Gemes / Simon May (Eds.), *Nietzsche on Freedom and Autonomy*, New

since we have no actual representation of a drive independent from the action, and so a drive cannot be construed independent from the action it is a basic unit for. Equally the action is in all cases what a person or an 'I' performs. The 'drives' and the 'I' should thus be construed as two *different* units of the same thing we call an 'act of will' from two *different* perspectives. The act of will for Nietzsche, which is borrowed from Schopenhauer, finds its representation in the action. This does not render the 'I' superfluous, it merely changes its role with respect to genuine agency and, for Nietzsche, debunks an age-old assumption. How we analyse the will and so how we determine what kind of a will it is, i.e. strong or weak, is no longer done on the basis of 'I' claims alone, but on how they relate to their corresponding actions. Genuine agency for Nietzsche is having a strong will, and thus a consistent relationship between one's 'I' claims and one's actions. The 'I' refers to the self as a whole, but it does not *determine* its unity or disunity, which is, for Nietzsche, determined by the relationship between it and its corresponding action. Notice also that the same is true of the rank order of drives: this order too merely refers to the agent or self as a whole, without claiming that the relation between the drives in the order is necessarily unified, i.e. weak or strong. Consequently, the 'I' and the drives refer to the same thing: the act of will, or better yet, the activity of willing or resolving to do something, which, for Nietzsche, is all that we are. They are therefore best construed as two *distinct* perspectives on this activity that employ two *distinct* concepts that merely characterize this distinction in perspective—nothing more.

There is a bridge between the two perspectives, for Nietzsche, and it is found in willing a course of action, which, when analysed, becomes apparent as an activity present and discernible in both perspectives. The activity of willing that we call the bridge, shows us that he is not committed to construing an action as a modification of a unified subject that passes on its inherent unity, but as the result of an activity that can, and often does result in unified action, but *not necessarily*. His sovereign individual's activity of willing *must* yield unified action because it is his model for genuine agency.⁵⁷ The key point for Nietzsche is that unified action is not solely a consequence of the inherent unity of the 'I' and practical agency, nor solely of the complex of drives. It is the consequence of an activity that can be described both in 'I'-terms and in drive-terms. The key philosophical point is that the first-person and third-person are merely two different perspectives on an activity, and neither can claim epistemic or

York 2009, pp. 89-106; Richardson, Nietzsche's Freedoms; and Christa Davis-Acampora, *Contesting Nietzsche*, Chicago 2013, p. 190.

⁵⁷ For an informative account and discussion on the sovereign individual as Nietzsche's attempt at a normative account of agency see Anderson, *What is a Nietzschean Self?*, pp. 228-232.

metaphysical superiority over the other with respect to agency without reduction, elimination or some kind of phantom.

5. Conclusion

The transcendental reading may have rescued first-person practical agency, and showed us a lack of fit between two Nietzschean perspectives on agency, but not resolved it. It moved the burden to a distinction between different types of practical agency. It was justified in arguing that *if* Nietzsche's aim is to eliminate first-person practical agency, then he must answer to its charges. The reading is wrong to charge Nietzsche with eliminativism, however. His attempt to naturalise agency is a critique on our conception of agency on the basis of the errors of self-knowledge. He rejects the idea that agency *as such* is exhausted by use of the first-person 'I', and by the claim that the self-conscious 'I' is separated from our experiences and actions.⁵⁸

We have argued that Nietzsche is aiming for a different approach to genuine agency, which does not leave us with construing the 'I' as a phantom, nor with elimination of first-person practical agency. But, in order to accommodate this approach, Nietzsche requires first-person practical agency. Correspondingly, in order to explain how practical agency fits with his rejections of the role of the unity of the 'I', he also requires what we call the bridge. If the bridge proves to be successful, then Nietzsche can coherently argue for a conception of agency as an activity, which is construed in first-person terms as *my* willing a course of action (*not* necessarily *my* following through), and in the third-person as the result of hierarchical drive-relations. The content of the 'I' (what 'we' attribute to ourselves) and the dispositional manifold (the drives) are thus the same thing construed from different perspectives. They are not two distinct things. The key to recognizing the bridge, we have argued, is to shift our

⁵⁸ A passage that is illuminating regarding Nietzsche's approach to consciousness is GS 354; also D 26 and 212. Another passage is In Z I, The Despisers of the Body: "Behind your thoughts and feelings, my brother, stands a powerful commander, an unknown wise man – he is called self. He lives in your body, he is your body". In these passages Nietzsche advances an account of consciousness that is subordinate to the organism and its development, not something transcendent, nor even transcendental to it. This entails that whatever consciousness adds to the drives, it must be explicable in their terms or subordinate to them. Think of the previous in terms of a probative question he poses: "How could anything originate out of its opposite?" (BGE 2) and he answers, "we can doubt, first, whether opposites even exist". In accordance with his own thoughts, we claim, it is misleading to construe consciousness and the drives as two things distinct in kind that interact. To construe consciousness and the drives as interacting and being mutually efficacious does not resolve the problem; it unsuccessfully sidesteps it.

focus from the relation between ‘I’-thoughts and drives, to that between ‘I’-thoughts and actions; or, if we use the drive as our basic unit of resolve, then we should shift away from the relation between drives and conscious thought to the relationship between drives and actions.

Genuine agency, for Nietzsche, is thus an individual’s willing or resolving a course of action and *simply* following through.⁵⁹

⁵⁹ We would like to thank Boris Demarest, Thijs Festjens, Kris Goffin, Annelies Monseré, Sarah Willems and two anonymous referees for their helpful comments and suggestions.