

University of Tartu
Institute of Philosophy and Semiotics
Department of Philosophy

On Personal Identity: the Epistemic Problem, Personal
Pronoun Revisionism, and *De Se* Attitudes

Master's Thesis in Philosophy

Litman Huang

Written under the supervision of
Prof. Dr. Daniel Cohnitz
Submitted on
14th August 2017

Table of Contents

1. Overview: Animalism & the Thinking Animal Problem	6
2. The Epistemic Problem	17
3. Response to the Epistemic Problem: Personal Pronoun Revisionism / the Linguistic Solution	20
4. More on Personal Pronoun Revisionism	24
4.1 Initial Analysis	25
4.2 Propositions as Fregean Sense and Reference	27
4.3 Fregean Analysis: <i>de se</i> Knowledge	32
4.4 Some Problems Concerning the Fregean Analysis	34
4.5 Return to the Epistemic Problem	37
5. Against Personal Pronoun Revisionism	38
5.1 Setting the Scene	40
5.2 The Argument	43
5.3 The Received View on <i>De Se</i> Attitudes	46
5.4 Defending The Argument Against RR	53
5.5 Defending the Argument Against CI	53
5.6 Possible Objections	55
6. Conclusion	59
Appendix: Noonan's Psychological Continuity Account	59
Reference	64

Abstract

The thesis focuses on the contemporary ‘animalism vs. neo-Lockeanism’ debate on personal identity over time. I will look at Noonan’s defence on behalf of neo-Lockeanism against the version of animalism formulated and popularised by Olson (1997, 2003, 2007). Noonan argues for neo-Lockeanism (1989/2003, 2010a, 1998) by rejecting the animalists’ *thinking animal argument*. In doing so, the neo-Lockean theorist commits to claims that give rise to the *epistemic problem*: how do I know which thinker is me, given the neo-Lockean claim that the thinking animal and the Lockean person are numerically distinct, but qualitatively identical entities capable of occupying the same temporal-spatial location? In response, Noonan devises *personal pronoun revisionism* (or the *linguistic solution*) as an attempt to resolve the problem: personal pronouns don’t refer to *the thinker* who thinks the I-thought e.g. the thought that ‘I am a person’, but instead they refer to *the person* who thinks the I-thought. According to the neo-Lockean view, since i) you know that *you* are a person viz. ‘you’ refers to a person and ii) persons are Lockean persons you can know that the thinker who is a person, the Lockean person, is *you* viz. ‘you’ always refers to the Lockean person not the thinking animal. Therefore, the *epistemic problem* is resolved since the question answers itself in the proposed semantic rule for the reference of personal pronouns. I argue that for considering problems of *de se* attitudes—propositional attitudes (*knowing that, thinking that, believing that* etc.) held towards propositions/beliefs identified by I-utterances such as ‘I am hungry’, ‘I want to have some ice-cream’ etc., Noonan’s response to the *epistemic problem* is not adequately motivated. With the *epistemic problem* unsolved, we have strong reasons for refuting neo-Lockeanism.

This thesis will have the following structure: In Section 1 I discuss Olson’s animalism as a position on personal identity. I will explain how Noonan’s refutation of the *thinking animal argument* leads to the *epistemic problem* as part of the broader *thinking animal problem*. Section 2 and 3 discuss this problem and Noonan’s proposed solution, *personal pronoun revisionism*, to it. Section 4 is an extended exposition of Noonan’s proposal. In Section 5 I argue that Noonan’s proposal fails for considering the problem of *de se* attitudes. Finally, I conclude

that with the problems Noonan's account faces, we have strong reasons to reject Noonan's neo-Lockeanism.

Abstract (Estonian)

Magistritöö teema on isiku identiteet ajas ja selle kohta käiv kaasaegne "animalism versus neolockeanism" debatt. Uurimise alla võetakse Noonani neolockeanismi pooldav argument Olsoni (1997, 2003, 2007) poolt formuleeritud ja populariseeritud animalismi versioon vastu. Noonan kaitseb neolockeanismi (1989/2003, 2010a, 1998), hüljates animalistide nn *mõtleva looma argumendi*. Nii tehes kirjutab see neolockeist alla argumentidele, mis on pannud aluse nn *episteemilisele probleemile*: kuidas saab mõtleja identiteedi kindlaks määrata, kui neolockeistide järgi on "mõtlev loom" ja "Locke'i isik" arvuliselt erinevad, kuid kvalitatiivselt identsed entiteedid, kes võivad eksisteerida samas ajalis-ruumilises asukohas. Vastusena loob Noonan nn *isikulise asesõna revisionismi* (või *lingvistilise lahenduse*) kui katse see probleem lahendada. Isikulised asesõnad ei viita *mõtlejale* (*the thinker*), kes mõtleb nn *mina-mõtte* (*I-thought*) näit mõtte "mina olen isik", vaid need viitavad hoopis *isikule* (*the person*), kes mõtleb *mina-mõtte*. Neolockeistliku vaate järgi, kuna i) sa ei tea, et *sina* oled isik viz. "sina" viitab *isikule* ja ii) isikud on Locke'i isikd, seega saad teada, et mõtleja, kes on isik (Locke'i isik), on *sina* viz. "sina" viitab alati Locke'i isikule, mitte *mõtlevale loomale*. Niisiis on episteemiline probleem lahendatud, sest küsimus vastab iseenelele ettepanud semantilises reeglis. Väidan, et *de se suhtumise* (st *propositsiooniline suhtumine* (teadma, mõtlema, uskuma etc), mis on suunatud mina-ütlustele nagu "olen nälgane") probleemide vaatluse alla võtmiseks ei ole Noonani vastus *episteemilisele probleemile* piisavalt põhjendatud. Kuna episteemiline probleem on lahendamata, on meil tungivalt põhjust neolockeanism kõrvale lükata.

Magistritöö ülesehitus on järgmine: Osas 1 kirjeldan Olsoni animalismi kui seisukohtavõttu isiku identiteedi suhtes. Seletan, kuidas Noonani vastuväide nn *mõtleva looma argumendile* viib välja (osana laiemast *mõtleva looma probleemist*) *episteemilise probleemini*. Osades 2 ja 3 vaatlen seda probleemi ja sellele Noonani poolt pakutud lahendust, nn *isikulise asesõna revisionismi*. Osa 4 on Noonani väite jätkatud seletus. Section 4 is an extended exposition of Noonan's proposal.

Osas 5 väidan, et Noonani ettepanekul ei õnnestu lahendada *de se suhtumise* (*de se attitudes*) probleemi. Kokkuvõttes järeldan, et võttes arvesse

Noonani väidetega kaasnevaid probleeme, on meil tugev põhjus hüljata Noonani neolockeism.

1. Overview: Animalism & the Thinking Animal Problem

The *epistemic problem*, which is part of the *thinking animal problem*, was a problem posed in response to the opponents of animalism, a view on personal identity formulated and popularised by Olson (1997:106-09, 2003:325-30, 2007:29-39). As a proponent of the neo-Lockean account on personal identity (1989/2003, 2010a,1998) Noonan addresses the problem posed by the animalist. In response, he develops so-called '*personal pronoun revisionism*' by (Olson 2007) or the '*linguistic solution*' by (Olson 2002). In this section, I will explain the *thinking animal problem* and how *personal pronoun revisionism* is supposed to (re)solve the problem. However, to get to the heart of the debate between Olson and Noonan, we first need to understand their respective views. I shall begin by presenting Olson's position, animalism, and the argument for it. It will soon be clear why the rejection of it will lead to the *thinking animal problem*.

Olson's animalism is backed by the *thinking animal argument*. It is a version of the biological criterion of personal identity, the tenant of which can be summarised in the slogan: *we are animals*.

The thinking animal argument

"It follows from these apparently trite observations that *you* are an animal. In a nutshell, the argument is this: (1) There is a human animal sitting in *your* chair. (2) The human animal sitting in *your* chair is thinking. (If you like, every human animal sitting there is thinking.) (3) *You* are the thinking being sitting in *your* chair. The one and only thinking being sitting in *your* chair is none other than *you*. Hence, *you* are that animal. That animal is *you*. And there is nothing special about *you*: *we* are all animals.

1. $(\exists x)(x \text{ is a human animal} \ \& \ x \text{ is sitting in your chair})$
2. $(x)((x \text{ is a human animal} \ \& \ x \text{ is sitting in your chair}) \rightarrow x \text{ is thinking})$
3. $(x)((x \text{ is thinking} \ \& \ x \text{ is sitting in your chair}) \rightarrow x = \text{you})$
4. $(\exists x)(x \text{ is a human animal} \ \& \ x = \text{you})$ "

(Olson 2003, p. 326)

Let the label 'animalism' can be understood as such. The defining feature of human animals is that they persist by means of preserving certain biological

properties e.g. the biological structure and the organs that are essential to the function of the central nerve system. If the conclusion, we are animals, is true, it follows that the persistence conditions of us are, like human animals, purely biological. Based on this animalist claim, Olson further claims that psychological connections are neither sufficient nor necessary for the persistence of us: i) we can survive total amnesia; ii) the preservation of psychological contents does not suffice for our survival in the future (1997: 107). Animalism defined as such falls under ‘the biological approach’¹.

Let us turn to the rival view to animalism, neo-Lockeanism, according to which the persistence conditions of persons² are purely psychological. In other words, it is sufficient for a person to persist over time if certain psychological connections or traits are preserved. This is a position under the ‘psychological approach’³.

The two views are ultimately incompatible. Neo-Lockeanism identifies persons as ‘psychological continuers’ and animalism identifies persons (us) as ‘biological continuers’. In virtue of having different persistence conditions, psychological continuers are *conceptually distinct* from biological continuers. Furthermore, there are cases in which they *actually* come apart in terms of occupying different spatial-temporal location. For example, human animals have foetus stages in which no mental activity is considered present. Presumably, embryos will develop over time up to the point that mental activities are undeniably said to exhibit. This example shows that the biological continuer *in fact* comes into existence *earlier* than the psychological continuer. Suppose the talk of temporarily coincident entities, we may say that they coincide for a certain period of time, and,

¹ Any attempt that analyses personal identity in terms of pure biological facts is defined as an account under the ‘biological Approach’. It shall not be confused with the ‘bodily Approach’ since an animal’s body is different than an animal: when at death, the animal’s corpse is qualified as ‘the body’ or ‘the animal’s body’ but not ‘the animal’. In such a case, we may say the animal *becomes* the corpse, instead of saying the animal *is* the corpse. The introduction of the biological approach and bodily approach is meant to respect just this difference. Under the bodily approach defined as such it is right to identify the person as the dead animal body or as the corpse, but under the biological approach a person is identified as an animal, something that has, if not essentially so, certain dynamic biological functions, but not identified as the body or as the corpse.

² I use “person” and personal pronouns such as “I”, “we” as synonyms here. Noonan takes it as trivially true (2010c). As revealed in Section 6, this is not as obvious as Noonan takes it.

³ Broadly speaking, any attempt that analyses personal identity in terms of pure psychological facts is defined as an account under the psychological approach.

very often, the psychological continuer will go out of existence first. Due to the existence of these cases in which they don't coincide permanently viz. coming into and going out of existence at the same time, it is a *contingent* fact that they could become *actually distinct* beings: *temporally distinct* in the above mentioned example, *spatially distinct* in the *prince/cobbler case* by Locke (see below).

While the existence of human animals as biological organisms is granted by common sense, the existence of psychological continuers or Lockean persons needs further support. This is typically done by thought experiments. A classic case for this is the *prince/cobbler case* by John Locke⁴.

- **Prince/Cobbler case**

Imagine there are two persons, the prince and the cobbler. The memories of the prince will be transferred to the cobbler's body, and subsequently the cobbler-body person will have all and only the memories of the prince's. That is to say, the cobbler-body person will remember himself as the prince. Now when the memory transfer process has finished, who is the person in the shoe of the cobbler's body, the prince or the cobbler?

Locke subscribes to the psychological approach, and argues that the cobbler-body person *is* the prince. He thinks that the later cobbler-body person with the prince's memories is *sufficient* for being identified as the prince, albeit he has 'switched' his body. The criterion for personal identity that underlines Locke's treatment of the case is the *memory criterion*. It implies that there is an entity *x* such that it persists in virtue of psychological connections, who is located where the prince's body is before the memory transfer and then where the cobbler's body is. For historical reasons, let us refer to entities of this sort as 'psychological continuers' or 'Lockean persons'.

The animalist must reject accounts of the psychological approach due to his adherence to the biological approach. In the *prince/cobbler case*, the cobbler-body is clearly not biologically connected with the prince-body person. The animalist theory does not allow the identification of the later cobbler-body person to the

⁴ Locke(1694/1975), chapter XXVII, "Of Identity and Diversity"

earlier person, the prince. In fact, the animalist must deny such identification. *Prima facie*, the *prince/cobbler case* is a case showing the existence of Lockean persons as psychological continuers because we have the intuition that persons *are* psychological continuers regarding cases that are structurally the same. Whether the Lockean person or the human animal (=the prince's body) is *the* person we call 'the prince' sums up the heart of the debate between the *biological approach* and the *psychological approach*.

I shall use the terms 'biological continuers' and 'human animals' to refer to entities such that their persistence conditions are purely biological; and 'psychological continuers' and 'Lockean persons' to refer to entities such that their persistence conditions are purely psychological. 'Biological continuer' and 'psychological continuer' are specially reserved for discourses that are explicitly about diachronic identity. In the *prince/cobbler case*, we say that there is a psychological continuer who is located where the prince's body is *before* the memory transfer takes place and *then* stays in the cobbler's body; and there is also a biological continuer who is located in the prince's body *before* and *after* the memory transfer (presumably, another biological continuer for the cobbler's body as well). Locke, in this case, identifies the psychological continuer as the person to whom we refer by the name "the prince", and therefore the biological continuer (what we may call 'the body') is ruled out as a person by his *memory criterion*.

- **Persistence conditions**

Biological continuers: purely biological

Psychological continuers: purely psychological

Moving back to the debate between Olson and Noonan, the animalist disagrees with the neo-Lockean theorist on *whether persons are biological continuers or psychological continuers* (it is important to note that this is, however, not an accurate presentation of Olson's position because it neglects i) subtle differences between personal pronouns, and the terms 'person' and 'people' that Olson draws; ii) Noonan agrees that human animals could have been psychological continuers. But for my present purposes, I shall not complicate the

matter, and proceed with the above framing). With this granted, we can formulate their views as follows:

- **Animalism**

All persons are biological continuers/human animals

- **Neo-Lockeanism**

All persons are psychological continuers/Lockean persons

Given animalism, it follows that at least *some of us* are indeed human animals. Therefore, the neo-Lockean claim that *all* persons are psychological continuers is false. To defend his position, Noonan must seek a way to resist the *thinking animal argument* that supports the animalist claim.

However, refuting the *thinking animal argument* is shown to be no easy task since none of its premises seems obviously problematic. Furthermore, denying either one of them would lead the neo-Lockean theorist to commit certain counter-intuitive claims, which, in turn, reveals inadequacy of neo-Lockeanism in reconciling our common-sense beliefs with philosophical discoveries as compared to its rival view animalism.

For example, premise (1) asserts the existence of a human animal that has certain biological features. Unless some radical view on the external world can be granted⁵, it is hard to see how one can deny the existence of such human animal claimed by premise (1). Premise (2) says that such animal can think. This can be disputed, however. Shoemaker thinks this is false⁶.

⁵ An idealist, for instance, would reject the existence of the thinking animal as a material object that exists in the external world.

⁶ On his view (1999a,1999b) human animals cannot think because the physical properties that determine the mentality of a person are only realised in the person, although the person, the human animal, and the coincident body (corpse-to-be) are all processors of such physical properties. What underlines this conclusion is i) the worry that the animalist may also face the *epistemic problem* (see section 2 below) for consideration of the corpse (the *corpse problem*, discussed in Olson 2004), ii) the metaphysics of properties and the nature of realisation Shoemaker independently defends (Shoemaker 1984, 1998). So the version of neo-Lockeanism Shoemaker endorses differs from that Noonan endorses. While Shoemaker's account and the underlying *corpse problem* surely worth of further investigation by their light, I would instead focus on the debate between Olson and Noonan. I shall proceed with the rather intuitive claim that

- **Thinking animal assumption**

Animals the persistence conditions of which are purely biological have the capacity to think

Noonan, unlike Shoemaker, endorses the *thinking animal assumption* above⁷. Since both Noonan and Olson agree that premise (2) is true, it is not the premise to fault at⁸.

Noonan denies the identification that follows from premise (3), and claims the following: “I am a thinking being and, of course, the thinking animal too, but *I am not that animal*”. Quite obviously, Noonan does not operate with the same ontology that the *thinking animal argument* supposes. The intent of the argument is that the thinking animal seems to be the *only* being at the location where you believe you are in, and thereby “you are that animal and that animal is you”. According to the animalist, Noonan’s reply will lead to the *thinking animal problem*, which consists of the following apparently highly counter-intuitive implications.

- **INCOHERENCE:** The neo-Lockean theorist’s acceptance of the *thinking animal assumption* and his subscription to the psychological approach to personal identity are incoherent. Given *thinking animal assumption*, a thinking animal must be a person after all in virtue of the fact that it can think just like

human animals are capable of thinking granted, and set aside the *corpse problem* from the main discussion.

⁷ (2010c, p.94) “(4) All normal healthy adult human animals are thinkers of true firstperson thoughts. This is what the animalist urges, rightly, I think, on the neo-Lockean as an evident common-sense truth”

⁸ By that I don’t mean premise (2) is true, and it should be reminded that Shoemaker (1999a, 1999b) offer arguments for thinking that it is false. Initially, the worry stems from the *corpse problem*: “Olson rightly ridicules the suggestion that the fetus goes out of existence when the person comes into existence. But it seems equally ridiculous to say that the corpse is something that comes into existence at death” (1999a, p.499) Hence, it seems that the body i.e. corpse-to-be is something that exists before my death and is such that it coincides with the person and the human animal. If this is so, then the body (corpse-to-be) must share the physical properties that the human animal (the person) has. If the realization of those physical properties will determine the mentality of the human animal, they must be also realised in the body (corpse-to-be) as well. Hence, we have *too many minds problem* i.e. the *epistemic problem* for the animalist himself. Since my focus is on the neo-Lockean claim that the Lockean person and the human animal are coincident *thinkers*, which I shall argue against in section 5, I shall at this moment suppose the truth of premise (2) and proceed.

the Lockean person does. If the Lockean person satisfies a psychological criterion and thereby is a person, so does the thinking animal. However, human animals are, after all, not Lockean persons because they differ in their persistence conditions. “[S]o the neo-Lockean’s attempt to identify the persistence conditions for persons collapses into incoherence, since he has to acknowledge different kinds of person with different persistence conditions – as if it were person-persons and animal-persons.” (Noonan 2010c, p.93)

- **MULTIPLICITY (Too Many Thinkers case):** Given premise (1) and (2), the neo-Lockean theorist would endorse that there is indeed a thinking animal located at where he is. But, the neo-Lockean theorist would not identify himself as the thinking animal because he is a person and that a thinking animal is not a person (human animals can at best be animal-person on pain of incoherence). It entails that there *also* exists an entity that *coincides* with the thinking animal occupying the same spatial-temporal location. This is this entity that the neo-Lockean theorist identifies persons with: Lockean persons. The neo-Lockean theorist would say that it is the Lockean person that is the *person* with whom he is numerically identical. Therefore, neo-Lockeanism entails multiplicity and the sharing of thoughts: there are two thinking *beings*, the thinking animal and the Lockean person, occupying the same spatial-temporal location and thinking the same thoughts.
- **MYTH:** Quantifying the coincident period, the thinking animal should have whatever cognitive capacities that the Lockean person can be said to have. Its thoughts have whatever complexity and sophistication the Lockean person has. In fact, during their coincident period, the thinking animal and the Lockean person have exact same thoughts, beliefs, desires, traits etc. Yet, the neo-Lockean theorist claims that they are *distinct* coincident entities. How can we make sense of this inflated ontology with this qualification on the capacities that the Lockean person and the thinking animal both possess?

The above three challenges encompass concerns from various

philosophical angles, namely: “(epistemic) how do I know that I am the Lockean person but not the human animal sitting here entertaining a thought shared with the Locekan person?”, “(metaphysical) how can we make sense of the claim that two numerically distinct beings can occupy the same space?”, “(ontological) what is the relationship between the Lockean person and the human animal with whom it coincides?”, “(semantic) which one of them does the personal pronoun “I” in a sentence “I am sitting” uttered by the Lockean person (or by the thinking animal) refer to?” These are difficult questions that make up the so-called *thinking animal problem* because they have the same root: they arise from the denial of the *thinking animal argument*. The *thinking animal problem* motivates the refutation of neo-Lockeanism by *reductio ad absurdum*:

Argument by reductio against neo-Lockeanism

1. If neo-Lockeanism is true, the *thinking animal problem* follows
2. The claims entailed by the *thinking animal problem* are absurd
3. Therefore, neo-Lockeanism is false

The neo-Lockean theorist must account for the concerns under the *thinking animal problem* and explain them to satisfactory degrees so as to resist the refutation of neo-Lockeanism by the above argument by *reductio ad absurdum*.

I shall now give an overview of Noonan’s responses on behalf of neo-Lockeanism.

- **INCORHERENCE:** Noonan rejects the animalist inference that due to its *present* sophistication in thoughts, the thinking animal is thereby a person. Noonan argues that he is not committed to the incoherency charged by Olson. He draws attention to the different *persistence conditions* between the thinking animal as a biological continuer and the Lockean person as a psychological continuer. In his view, to say that an entity x is a person is to say something about x’s *entire history* including x’s past, present, and future⁹. The thinking animal as a biological continuer has different history than the Lockean person

⁹ Noonan 2010c p.97

as a psychological continuer has: they, at least, as a matter of contingent fact, differ from their past histories, for the thinking animal as a biological continuer has a foetal stage, but the Lockean person as a psychological continuer doesn't. The substantial neo-Lockean claim is the following *de dicto* necessary truth: if x is a psychological continuer, x is a person. The period of existence being quantified in the discussion includes the coincident *and* non-coincident periods. Given this, it entails that the thinking animal does not satisfy the neo-Lockean *de dicto* principle above, which governs the concept of persons. Even if the thinking animal as a biological continuer does satisfy the psychological criterion for certain period (the period when the Lockean person as a psychological continuer also exists), in virtue of having different persistence conditions the animal is *not*, as a matter of *de dicto* truth, a psychological continuer.

- **MULTIPLICITY:** Multiplicity is accepted by Noonan¹⁰. However, while it is true that during the coincident period, the thinking animal and the Lockean person occupy the same spatial-temporal location, it is not true if we quantify over their *entire* histories. That is to say, they don't occupy *exact* same *temporal* location. Noonan defends the four-dimensionalist thesis that persons are *perdurants*¹¹: persons are objects that comprise of maximal temporal aggregates of person-stages (temporal parts of persons). This commitment to the four-dimensional ontology allows him to explain the multiplicity as follows: there are some temporal stages such that they are shared by the thinking animal as a biological continuer and the Lockean person as a psychological continuer. MULTIPLICITY would be difficult to accept if the claim is that permanently coincident beings are still numerically distinct beings. But this is not Noonan's claim, and the contingent facts concerning the sequence of their coming into and going out of existence tell us that we are not dealing with permanently coincide entities. Noonan argues that this is not an *ad hoc* response to the *thinking animal problem* since four-dimensionalism can be defended

¹⁰ Noonan 2012 p.317

¹¹ Noonan 1989/2003 p. 118-124

independent of the discussion of personal identity over time¹².

- **MYTH:** Noonan claims that the temporarily coincident thinking animal and Lockean person are cognitively identical i.e. qualitatively indistinguishable in their cognitive behaviours¹³. Nevertheless, they are *numerically distinct* entities due to the fact that they have different persistence conditions and the above-mentioned contingent fact they are in fact not permanently coincided¹⁴. What we have is a case in which two thinking beings share the same thoughts.

As the responses indicate, INCOHERENCE is related to the discussion of discovering persistence conditions for persons. This is related to Noonan's view on thought experiments and the structure of the concept of personal identity. MULTIPLICITY turns on the ontological question whether persons persist as *endurers* or *perdureres* i.e. whether persons have temporal parts. This corresponds to Noonan's discussion of the possibility of indeterminacy in personal identity, the consideration for which leads to the four-dimensional view on the persistence of persons¹⁵. Finally, MYTH leads to the specific thesis *personal pronoun revisionism* devised to respond to the *epistemic problem*. In this thesis, I wish to focus only on MYTH (see appendix for some details about Noonan's defences for the concept of personal identity and four-dimensionalism).

One of the criticisms Olson (2007) makes to Noonan's response to the MYTH challenge is that it creates an irresolvable epistemic problem (hereafter *the epistemic problem*): given MULTIPLICITY and the coincident thinkers being cognitively identical, how do I know which of the thinkers is me? For all I know, I can comprehend I-utterances, being able to express I-thoughts, and I am a person. But following the neo-Lockean view, there also exists a *distinct* human

¹² Noonan 1989/2003 p.118-124

¹³ Noonan 2012, p.317

¹⁴ Noonan 2010c p.93 note 4 thinks that permanently coincident objects are identical. "I think the permanently coincident objects are identical and I do not think that merely future differences can determine whether something is now a thinker.

¹⁵ "The key point is that the only way to make comprehensible the idea that statements of identity can be indeterminate in truth-value is by appeal to the idea of semantic indecision. But in order to apply this idea to the explanation of indeterminacy in statements of identity over time one has to acknowledge ontological commitments which fit more comfortably with the four-dimensional theorist's framework than with its rejection" Noonan 1989/2003 p.118

animal as a biological continuer with which my thoughts are shared. The thinking animal can think just like me. How do I know if I am the person thinking truly but not the animal thinking falsely that it is a person?

The *epistemic problem* and Noonan's response will be the focus of our attention in the remaining of the discussion. I shall first explain in more detail what is at issue, and then critically examine Noonan's *personal pronoun revisionism / linguistic solution* (I will speak of them interchangeably) as a follow-up response to the problem. Ultimately, I argue that Noonan's proposal, despite its ingenuity, fails for considerations of the problems of *de se* attitudes.

2. The Epistemic Problem

In this section I will take a closer look at *the epistemic problem* introduced in the last section. What underlines the *epistemic problem* is Noonan's acceptance of MULTIPLICITY viz. there are two numerically distinct but qualitatively identical entities that "co-inhabit" for a certain period of time the same spatial location, a Lockean person and a human animal. MULTIPLICITY a.k.a. the *too many thinkers* case follows from the (neo-Lockean theorist's) commitments below:

1. The persistence conditions of human animals (biological continuers) are purely biological;
2. The persistence conditions of Lockean persons (psychological continuers) are purely psychological;
3. Lockean persons can think;
4. Human animals can think;
5. In normal cases, each Lockean person coincides with a human animal for a certain period such that the Lockean person as a psychological continuer comes into existence later than the human animal as a biological continuer (presumably, the psychological continuer will go out of existence before the biological continuer)

(1) is backed by the common sense truth that there are human animals which can exist over time; (2) is backed by cases structurally the same as the

prince/cobbler case by Locke. (3) follows from (2), and (4) is the *thinking animal assumption* agreed mutually by Noonan and Olson. (1) to (4) leads to (5): MULTIPLICITY / the *too many thinkers case*. But MULTIPLICITY by itself is neutral to neo-Lockeanism and animalism, for it does not tell us whether persons are thinking animals or Lockean persons. The addition of the thesis: *all persons are psychological continuers* completes Noonan's neo-Lockeanism (again, see appendix for details).

From (5), (3), and (4) (and Noonan's materialist view on persons, see CI in section 6.1), we can infer that the thinking animal is a *cognitive peer* with the Lockean person during their period of coincidence in the sense that they are *cognitively identical* thinkers sharing the same thoughts. Let us employ the label '*too many thinkers case*' for the MULTIPLICITY case as such:

- **Too many thinkers case**

On the neo-Lockean view, whenever a Lockean person thinks, the "co-inhabited" human animal thinks in the Lockean person as well. They both entertain the same thoughts. They are thus in this respect *cognitively identical*. But, they are *numerically distinct* due to having different persistence conditions and the contingent fact concerning the existence of human animals and Lockean persons.

According to Olson, the *too many thinkers case* leads to the *epistemic problem*.

"...it is hard to see how you could ever know which thinker you are, the animal or the person (the one with psychological identity conditions): the epistemic problem"

"that thing [thinking human animal] would seem to be psychologically and epistemically indistinguishable from you, to the point of having the same beliefs about itself as you have about yourself, and the same grounds for them. (...) even though there are beings of the sort we take ourselves to be sitting in our chairs and thinking our thoughts, we can have no way of knowing whether we are those beings."

(Olson 2002 p.2, p.4)

This is the *epistemic problem*. The problem arises due to Noonan's acceptance of the existence of multiple thinkers and their being cognitively identical. If I am a person on the neo-Lockean view, there is a thinking animal that is "psychologically and epistemically indistinguishable from me". How can I know which one of the thinkers is me? This leads to a form of scepticism because many of my deep convictions such as beliefs about myself turn out to be ungrounded: I certainly have beliefs about myself and it seems to be trivially true that I am a person. But the temporally coincident thinking animal and Lockean person are cognitively identical. If, following Noonan, only the Lockean person as a psychological continuer is a person, how do I know that *I* am, in the sense of numerical identity, the Lockean person? For all I know, it could be, at least conceivably, that I am the animal thinking falsely that I am a person. If this is so, then many of the beliefs I have about myself will turn out to be false, even the belief 'I am a person' (for human animals are not persons). The *epistemic problem* leads to deep, if not outrages, scepticism, for it not only leaves open the question whether I am a person but also suggests that I *could not possibly* know the answer. What can I appeal to in distinguishing the person-thinker from the non-person thinker? Even thought metaphysically speaking it is true that one of them is a person and is numerically identical with me, how do I know?

- **The epistemic problem (too many thinkers problem)**

In the coincident case in which two thinking beings, the animal and the Lockean person, are thinking the exact same I-thoughts occupying the exact same spatial-temporal location, how do I know which one of the thinkers is me?

Note that this problem carries implications to the broader animalism vs. neo-Lockeanism debate. The *epistemic problem* is a consequence of the *too many thinkers case* as part of the neo-Lockean view, which seems to be highly counter-intuitive, if not by itself absurd. Unless the neo-Lockean theorist can offer a satisfactory explanation, the *argument by reductio against neo-Lockeanism* is motivated.

3. Response to the Epistemic Problem: Personal Pronoun Revisionism / the Linguistic Solution

Noonan's response to the *epistemic problem* is based on his proposed semantic rule for first-personal pronouns. In this section, I will give an overview of his account.

Let us begin with the distinction that is fundamental to his response: the distinction between *thinkers of I-thoughts* and *objects of first-person reference*.

“... what the psychological continuity theorist must do ... is reject the original definition of a person, substitute for it the notion of the object of self-reference, and insist on a *distinction between the 'I'-user and the reference of 'I'*. Then he can say that when a person and a human animal coincide the animal can indeed think 'I'-thoughts, but is not thereby constituted a person, *for the reference of its 'I'-thoughts is not itself but the person with whom it is sharing those thoughts.*” (Noonan 2010b, p. 197, my italic)

“The word 'person' as it is used in this debate (and its synonym 'self') is merely intended to allow a nonindexical formulation of these questions. So, 'person' in the philosophical debate simply means object of first-person reference.” (Noonan 2010c, p.95)

The first quote sums up such distinction. It is meant to contrast with what Noonan calls “the original definition of a person” which is, presumably, that x is a person iff x is a thinker of I-thoughts. Given the neo-Lockean view that all persons are Lockean persons / psychological continuers, it follows that while both the human animal and the Lockean person are indeed thinkers of the same I-thoughts, e.g. *that I am sitting*, the 'I' in the I-thought refers to the Lockean person instead of the thinking animal (hereafter, I shall speak of thoughts with the italic that-clause form: *that I am sitting*). In virtue of being a thinker of the thought *that I am sitting*, the thinking animal, according to Noonan, is not thereby constituted to be a person¹⁶.

¹⁶ So this proposed rule for the reference of personal pronouns serves the interest of the neo-Lockean theorist. But it is worth reminding that what really rules out human animals, albeit being capable of being thinkers of I-thoughts, is that i) the concept of persons is structured around a set of necessary and sufficient conditions ii) the *transplant intuition* we have regarding cases that are structurally the same as the *prince/cobbler case*, e.g. the *Brown/Brownson case* by Shoemaker (1963). See appendix for details.

In the second quote, it indicates Noonan's conviction that personal pronouns such as 'I', the word 'self', and the word 'person' are all referential expressions that are *on a par* regarding their meaning: these are expressions that have the same linguistic reference i.e. objects of first-person reference, although they differ in terms of some other linguistic aspects, namely, while 'I' is an indexical sensitive to contexts, 'person' is not. Noonan takes this difference as trivial in the debate of personal identity¹⁷.

Return to the *epistemic problem* the neo-Lockean theorist encounters. The quick answer by Noonan is that the question "answers itself".

"Hence, there is no epistemic worry since the question whether I am a person or an animal *answers itself*, and there is no worry that we must say that there are two persons present, since there is a difference between the person and the animal which justifies denying that the animal is a person, namely that it is not an object of first-person reference, a self" (Noonan 2012, p.317, my italic)

When x has an I-thought *that I am a person*, as the idea goes, the 'I' in x's I-thought refers to *the person who thinks the I-thought*. Suppose x is a person in the neo-Lockean sense. It follows that x's I-thought *that I am a person* is about herself. Hence, x can know that she is *the person who thinks the I-thought* (only one of the thinkers is a person) in virtue of being both the thinker and the object of the I-thought *that I am a person*. On the other hand, although the human animal y that coincides with x also knows *that I am a person* viz. the person who thinks the I-thought is a person, the same I-thought that y grasps is *not* about y. Both the animal y and the Lockean person x know *that I am sitting*, while the 'I' in the I-thought refers to the person who is thinking the thought i.e. the Lockean person in this case. They both know the same I-thought, but only the Lockean person x is thinking *of herself*.

If we take the animal y's perspective, reporting the animal's belief, we may say that the animal knows that *it* is a person, while the '*it*' is an indirect reflexive

¹⁷ "(1) Persons and only persons are objects of first-person reference. "I take this to be trivially analytic. If anything is a person, it is capable of being an object of (its own) first-person reference and anything which is capable of being an object of first-person reference is a person." (Noonan 2010c p.93-4)

term referring to the Lockean person x, not the animal y itself¹⁸. By contrast, the Lockean person knows that *it (s/he)* is a person, while the '*it (s/he)*' is a direct reflexive term because the Locekan person x is the object of the I-thought (for the Lockean person is a person). Presented in a non-indexical way, we say both the Lockean person x and the thinking animal y know *truly* the thought *that I am a person* viz. the person-thinker of the I-thought to which the 'I' in the I-thought refers is a person. No epistemic worry. Let us summarise the idea of this response in neutralised terms as follows:

- **Response to the epistemic problem**

For the same I-thought, there can be a person-thinker as well as non-person-thinker of such a thought. When a thinker x has an I-thought, x is thereby qualified as a thinker of the I-thought *that I am a person*, but this alone is not sufficient for x to be the object for the reference of the 'I' in the thought.

According to Noonan's proposal, the 'I' in x's I-thought refers to entities that are such that i) x is a person and ii) x is a thinker of the I-thought. If x is a person-thinker and x is has the thought *that I am a person* then x is the object for the reference of the 'I' in such I-thought; whereas if x is *not* a person-thinker and has the same thought, x's thought *that I am a person* is not about x, even though x is the thinker of it. Nevertheless, both of person-thinker and the coincident non-person-thinker know the I-thought *that I am a person* because they both know *that the person whose thinking the I-thought is a person* (but the thought only refers to the person-thinker).

If this explanation holds, then not only does it rescue neo-Lockeanism from *reducio ad absurdum*¹⁹, since it allows the neo-Locekan theorist to avoid the radical scepticism posed by the *epistemic problem*, but also it threatens animalism

¹⁸ "The animal, on the other hand, does not know of itself that it is a person, since it is not. It does, however, know the thought it would express by saying 'I am a person'. Does it know that it is itself a person? Is it correct to say 'the animal knows that it itself is a person'? Yes: 'it itself' is here the indirect reflexive. It is not an expression whose reference is the animal. In fact, it does not refer at all. 'The animal knows that it itself is a person' is a report that the animal knows what it could correctly express by saying 'I am a person'" 2010b, p. 198

¹⁹ In fact it also answers the Incoherence challenge and the personhood problem posed by Olson (2002)

in return: i) 'we' refers to persons²⁰; ii) persons are psychological continuers; iii) human animals are not psychological continuers; therefore, iv) we are not human animals. The last claim is in direct conflict with the animalist's tenant that we are animals.

Having said that, Noonan's proposal is not necessarily incompatible with animalism since it is possible to rephrase Noonan's proposal in neutral terms. The animalist *can* agree with Noonan that the *linguistic solution* can indeed resolve the *epistemic problem* [fn] This is especially desirable if the animalist accepts that animalists also face the *epistemic problem* due to the *corpse problem* as suggested in Shoemaker (1999a, 1999b)[/fn], but still argues against the built-in neo-Lockeanism claim—all persons are psychological continuers. After all, the requirement we found in the proposal is just that the objects of first-person reference are persons. The proposal by itself doesn't imply the above neo-Lockean claim. The animalist could accept the proposal *and* MULTIPLICITY, but argue that persons, contrary to Noonan, are human animals as biological continuers. If the discussion goes along this line, we move back to the initial disagreement that characterises the debate between animalism and neo-Lockeanism. The animalist may attempt to explain or explain away the *transplant intuition* as (Snowdon 1991, Olson 1997) did (see appendix for more details), a move anticipated and defended by Noonan²¹.

This is not how the discussion actually develops between Olson and Noonan. Olson (2002) argues that Noonan's proposal is not a satisfactory response to the *epistemic problem*. But before we get to the criticisms of Noonan's proposal, I think it is necessary to linger on Noonan's proposal so as to get down to the details.

²⁰ On the assumption that personal pronouns refer to the same objects as the term 'person' refers to.

²¹ "That is, why can quite different kinds of thing with quite different kinds of persistence condition not be objects of first-person reference (Noonan 1978: 351)? This is a very good question. The only answer, I think, is the transplant intuition, which has to be accommodated." (2010c, p.98)

4. More on Personal Pronoun Revisionism

In this section, I will focus on Noonan's further explanation for his *personal pronoun revisionism* as in Noonan (2010b, 2010c, 2012). I shall note ahead that the general tension we will see in his proposal is the contention that *the thinking animal cannot refer to itself when having an I-thought*. Let us begin with some of Olson's remarks on Noonan's proposal.

"The idea is that the epistemic problem arises only if we assume that any being that says or thinks 'I' thereby refers to itself. (...) Noonan says it is false.

(...) 'I' doesn't simply refer to whatever utters it. It refers to whatever person utters it. When a non-person says 'I', that pronoun refers to the person who shares in that utterance. (...) *And first person thoughts share their content with the first-person utterances that express them.* (...) The belief that one is a person and not an animal is guaranteed to be true, no matter what being has it."

(Olson 2002, p.8)

One sharp difference between what Noonan calls the "original definition of person" and his revisionism is that the reference of 'I' on the former view is the thinker of the I-thought, and the reference of 'I' on the latter view is the person who thinks the I-thought. What is being "revised" here is indicated by the extra '*personhood requirement*', and strictly speaking it is not really a revision but a further constraint in addition to being *the thinker of*. On the original view, the derived characterisation of the contents of I-thoughts has the attribute of being *reflexive*: if x is a thinker of the I-thought *that I am so-and-so*, the thought refers to x. On Noonan's view, due to the *personhood requirement* and neo-Lockeanism, contents are not necessarily reflexive in this sense, and in the coincident case it is not. For the animal is the thinker of the thought *that I am so-and-so* but not the object of the first person reference of the thought on the neo-Lockean view. Noonan remarks that it "leads to the surprising conclusion that there are more thinkers than common-sense acknowledges and that some of these are not objects of their first-person reference" (Noonan, 2010c, p.95).

So, we have a case in which though the thinking animal and the Lockean person are both thinkers of the same I-thought, the content of the thought is *about*

the person only—the Lockean person. This is the element in Noonan’s revisionism that seems to be difficult to take on board. It is because we usually think the contents of I-thoughts are essentially reflexive. This seems to be what we understand of and mean by I-utterances. The implication that the non-person-thinker, though being cognitively identical to the person-thinker, is not referring to itself while making I-utterances, but the person-thinker is, seems more than just being counter-intuitive. For, the proposal suggests ‘I’ being reflexive is not necessary. In the case of the coincident non-person-thinker, it is not.

Being aware of this tension Noonan attempts to explain away the counter-intuitiveness by the appeal to the Fregean approach to contents of thoughts. I shall devote the next sub-sections for scrutinising this idea.

4.1 Initial Analysis

We begin with his proposed rule for the reference of ‘I’.

“The rule of reference for ‘I’ is thus not that the reference of ‘I’ is the thinker of the ‘I’-thought (there is no unique thinker), but that *the reference is the person thinking the ‘I’-thought*. No other rule is needed ...” (Noonan 2010b, p.197, my italic)

The proposed rule of the reference of personal pronouns such as ‘I’ is: *the person thinking the I-thought*. For example, if person A has an I-thought *that I am sitting* and the corresponding I-utterance is “I am sitting” by A, then the ‘I’ in the I-utterance refers to the person who is thinking the thought *that I am sitting*, that is, A, and the thought *that I am sitting* is a thought about A.

- **Rule for the reference of ‘I’**

x is the reference of ‘I’ iff i) x is the thinker of the I-thought expressed by the I-utterance, and ii) x is a person

Applying this proposed rule to the coincident case, we say both the human animal and the Lockean person satisfy (i), but only the Lockean person satisfies (ii) on the neo-Lockean account. Thus, the reference of ‘I’ is the Lockean person. (i) and (ii) are the conditions for an entity x being the reference of a personal pronoun.

The following quote tells us how the proposed rule resolves the *epistemic problem*.

“Both the person and the animal can know that their utterance of ‘I am a person’ is true. *And each understands this sentence, so each knows the proposition it expresses (...)*. In virtue of knowing this, the animal knows, *of the person*, that it is a person, and the person (who, of course, knows no less) knows, *of the person*, that it is a person. So the person knows, *of himself*, that he is a person (if x knows *of x* that he is F, then x knows *of himself* that he is F). I know that what I express by ‘I am a person’ is true. *Since I understand this sentence I also know the proposition thus expressed*, and in doing so I know, *of myself*, that I am a person. But no more than this can be required for knowledge *de se*, of oneself, that one is a person.” (Noonan 2010b p.198, my italic)

The first thing to note is the acknowledgement that understanding an utterance/sentence amounts to *knowing* the proposition it expresses. In the course of making an I-utterance ‘I am a person’, both the human animal and the Lockean person understand this utterance and thereby they *know* the proposition it expresses. It is the proposition that is said to be the entity being known. More precisely, we hold the propositional attitude *knowing* towards the proposition expressed by an I-utterance. It is a relation between a knower and a proposition. It implies that Noonan subscribes to a certain theory of propositions when it comes to matters of philosophy of language. As we shall see shortly, it is a Fregean theory.

My second note is that there nevertheless is a difference between the animal’s knowing *that I am so-and-so* and the Lockean person’s knowing *that I am so-and-so*, despite the suggestion that their thoughts have the same contents viz. express the same proposition. Let P be the proposition that the I-utterance ‘I am so-and-so’ expresses. As Noonan suggests in the quote, there is a sense in which the Lockean person’s knowing that P is the same as the animal’s knowing that P captured under the gloss “they both know, of the person, that P”. But, most importantly, there is another sense in which their knowing that P are somewhat *different*: only the person knows, *of herself*, that P (and not both but only one of them is a person). Regarding this sense, on the neo-Lockean *view only* the

Lockean person who knows, of herself, that P. Stated free of indexicals, we may say that the Lockean person knows, of the Lockean person, that P, and the animal knows, of the Lockean person, that P. Such difference is marked as knowing *de se* / *de se* knowledge: x knows *de se* that P iff P is about x itself iff the subject of P is x.

4.2 Propositions as Fregean Sense and Reference

Propositions are analysed in terms of Fregean thoughts and senses.

“So the question for the first-person pronoun revisionist is not how the animal and person can differ in their cognitive capacities, given their coincidence, but why, given their numerical distinctness, they are cognitively identical, are entertaining exactly the same *Fregean thoughts*.” (2012, p.317-8, my italic)

“I find it helpful to think of this in Fregean terms, following Geach (1976). The first-person pronoun ‘I’ has a constant linguistic meaning but expresses different senses in the mouths of different utterers. The sense of ‘I’ as uttered by x is a function of x; it is a different sense from that expressed by y if x is a person distinct from the person y. We can speak of the ego function which carries any x to the sense ‘I’ has in x’s mouth. Then, allowing Greek letters as variables ranging over senses, we may say that ‘x thinks that it itself is F’ means ‘for some α , $\alpha = \text{ego}(x)$ and x thinks [α is F]’—here the thought that [α is F] is the thought you would express in language by attaching the predicate ‘is F’ to a subject term whose sense is the mode of presentation α .

So it is true that the animal thinks that it itself is a person. But it is not true that the animal thinks *of* itself that it is a person, since the assumption required for this inference is that in all cases ego (x) is a mode of presentation of x, and if x is not a person this is not so; in this case, ego (x) is a mode of presentation of the person with whom x shares the first-person thought.” (Noonan 2010b, p.198-9)

There are several important characterisations being laid down in the second quote. The most important ones are those related to the sense of ‘I’: i) the word ‘I’ expresses different senses relative to “the mouth of the utterer”; ii) a sense is a mode of presentation of some x; iii) there is an ego function of a person x to a sense of ‘I’; and lastly, iv) the word ‘I’ has constant linguistic meaning.

A Fregean thought is a complex of Fregean senses. Consider an I-utterance ‘I am hungry’ by person A. The ego function “carries person A to a sense

of 'I' in A's thought", and A's thought consists of such sense of 'I' and the sense expressed by the predicate '...am hungry'. Together, the sense of 'I' and the sense of '... am hungry' constitute A's Fregean thought, which is what we have been calling 'I-thought' described in the form: *that I am hungry*.

Example: suppose there are two persons A and B who both make the same utterance 'I am hungry'. The corresponding Fregean thought that person A has is:

($T^x_{\Phi=\text{ego}(y)}$ reads T: thought; x: thinker of T; Φ : sense of 'I' which is a mode of presentation; $\text{ego}(y)$: ego function of person y)

- A utters 'I am hungry'
- A thinks a Fregean thought $T^x_{\Phi=\text{ego}(y)}$ identified by A's utterance 'I am hungry'
- A thinks that $T^A_{\Phi=\text{ego}(y)} =$ the sense of 'I' + the sense of '... am hungry'
- A thinks that $T^A_{\Phi=\text{ego}(y)} =$ {the sense of 'I' = $\text{ego}(y)$ } + {the sense of '... am hungry' = [is hungry]}
- A thinks that $T^A_{\Phi=\text{ego}(y)} =$ the sense of 'I' = $\text{ego}(y)$ + [is hungry]
- A thinks that $T^A_{\Phi=\text{ego}(y)} =$ the sense of 'I' = $\text{ego}(A)$ + [is hungry]
- A thinks that $T^A_{\alpha=\text{ego}(A)} =$ α + [is hungry]
- A thinks that $T^A_{\alpha=\text{ego}(A)} =$ [α is hungry]
- A's Fregean thought = $T^A_{\alpha=\text{ego}(A)}$

In the case of person B, although B's utterance is the same as A's, but the corresponding Fregean thought that B has is different. It is because the sense of 'I' that constitutes B's thought differs from the sense of 'I' that constitutes A's thought in the *object* of which it is a mode of presentation.

- B utters 'I am hungry'
- B thinks a Fregean thought $T^x_{\Phi=\text{ego}(y)}$ identified by B's utterance 'I am hungry'

- B thinks that $T_{\Phi=\text{ego}(y)}^B =$ the sense of 'I' + the sense of '... am hungry'
- B thinks that $T_{\Phi=\text{ego}(y)}^B =$ {the sense of 'I' = ego(y)} + {the sense of '... am hungry' = [is hungry]}
- B thinks that $T_{\Phi=\text{ego}(y)}^B =$ the sense of 'I' = ego(y) + [is hungry]
- B thinks that $T_{\Phi=\text{ego}(y)}^B =$ the sense of 'I' = ego(B) + [is hungry]
- B thinks that $T_{\beta=\text{ego}(B)}^B =$ β + [is hungry]
- B thinks that $T_{\beta=\text{ego}(B)}^B =$ [β is hungry]
- B's Fregean thought = $T_{\beta=\text{ego}(B)}^B$

Due to different senses as constitutants, it follows that the same I-utterance by A and B expresses *different* Fregean thoughts they have respectively: $T_{\alpha=\text{ego}(A)}^A \neq T_{\beta=\text{ego}(B)}^B$.

Note that the *argument* of the ego function is supposed to be a person x, and the *value* of the ego function would be a sense Φ which is a mode of presentation of the very person x. In normal cases in which the utterer of an I-utterance is presumed to be a person, a one-to-one correspondence is exhibited between i) the argument of the ego function, ii) person x, and iii) the object of which the sense of 'I' Φ is a mode of presentation. In cases like A's and B's, we witness that the *thinker* of the Fregean thought, the argument of the ego function, and the object of which a sense of 'I' is a mode of presentation of are 'aligned': they are the *same* person x. Noonan's distinction between thinker of I-thought and objects of first-person reference is thus unnecessary regarding normal cases. It is because the thinker of the thought is also the referent of the thought e.g. A is the thinker of $T_{\alpha:\text{ego}(A)}^A$, and the sense of 'I' α is a mode of presentation of A.

Contrasting the analysis of the coincident case, matters stand differently. We would need to work under the following neo-Lockean terms: a) since all persons are psychological continuers, only the Lockean person can be the argument of the ego function; b) the Lockean person and the thinking animal are cognitively identical, so their Fregean thought(s) cannot differ. Due to (a) we can expect that the case for the Lockean person is basically the same as the normal

cases in which the thinkers are presumed to be persons, but different in the case of the coincident animal.

- (HA: human animal, LP: Lockean person)
- HA utters 'I am hungry'
- HA thinks a Fregean thought $T^x_{\phi=\text{ego}(y)}$ identified by HA's utterance 'I am hungry'
- HA thinks that $T^{\text{HA}}_{\phi=\text{ego}(y)}$ = the sense of 'I' + the sense of '... am hungry'
- HA thinks that $T^{\text{HA}}_{\phi=\text{ego}(y)}$ = {the sense of 'I' = ego(y)} + {the sense of '... am hungry' = [is hungry]}
- HA thinks that $T^{\text{HA}}_{\phi=\text{ego}(y)}$ = the sense of 'I' = ego(y) + [is hungry]
- HA thinks that $T^{\text{HA}}_{\phi=\text{ego}(y)}$ = the sense of 'I' = ego(LP) + [is hungry]
- HA thinks that $T^{\text{HA}}_{\gamma=\text{ego}(LP)}$ = γ + [is hungry]
- HA thinks that $T^{\text{HA}}_{\gamma=\text{ego}(LP)}$ = [γ is hungry]
- HA's Fregean thought = $T^{\text{HA}}_{\gamma=\text{ego}(LP)}$

Notice that the ego function takes the Lockean person (LP) instead of the thinking animal (HA) as its argument despite the fact that the thinking animal is, like the Lockean person, a thinker of the (/its own) Fregean thought. This is because the ego function takes only the *person* who utters 'I am hungry', and neo-Lockeanism implies that it is the Lockean person (it would be the thinking animal if the ego function is free of the *personhood requirement* and takes whoever utters the I-utterance i.e. $T^{\text{HA}}_{\phi=\text{ego}(\text{HA})}$ instead of $T^{\text{HA}}_{\gamma=\text{ego}(\text{LP})}$). As a result, the value of the ego function γ is a sense of 'I' that is a mode of presentation of the Lockean person. In short, the thinker of the Fregean thought, the argument of the ego function, and the object of which a sense of 'I' is a mode of presentation are *not* 'aligned', and the *personhood requirement* as a substantial element proposed by *personal pronoun revisionism* manifests itself as a condition on the ego function. Compare the case for the Lockean person:

- LP utters 'I am hungry'

- LP thinks a Fregean thought $T^x_{\phi=\text{ego}(y)}$ identified by LP's utterance 'I am hungry'
- (...)
- LP thinks that $T^{\text{LP}}_{\phi=\text{ego}(y)}$ = the sense of 'I' = **ego(LP)** + [is hungry]
- LP thinks that $T^{\text{LP}}_{\gamma=\text{ego}(\text{LP})}$ = **Y**+ [is hungry]
- LP thinks that $T^{\text{LP}}_{\gamma=\text{ego}(\text{LP})}$ = [**Y** is hungry]
- LP's Fregean thought = $T^{\text{LP}}_{\gamma=\text{ego}(\text{LP})}$

The correspondence between the thinker of the I-thought, the argument of the ego function and the object of which the sense of 'I' γ is a mode of presentation are exhibited: the same Lockean person LP. What determines the same Fregean thought ($T^{\text{HA}}_{\gamma=\text{ego}(\text{LP})} = T^{\text{LP}}_{\gamma=\text{ego}(\text{LP})}$) of the animal and the Lockean person is the same sense of 'I' γ . Take the perspective of the thinking animal for the moment. Since the sense of 'I' γ requires the object of which γ is a mode of presentation to be a person, it cannot be a mode of presentation of the thinking animal. γ , after all, must be a mode of presentation of some objects. Thanks to the *personhood requirement* proposed by Noonan, it follows that γ "is a mode of presentation of the person with whom the thinking animal shares the first-person thought": the Lockean person.

If we follow Noonan in analysing propositions as entities expressed by I-utterances in the above Fregean way, then we should expect that Fregean thoughts play the *cognitive role* that propositions are commonly expected to play in a theory of belief. If it can be said that I now have a thought and am ready to express what I know by the utterance 'I am a person', what I know is the Fregean thought identified by such utterance i.e. the proposition, the belief, expressed by my utterance 'I am a person'. Likewise, if I understand someone's I-utterance 'I am hungry', what I understand is the Fregean thought expressed by that person's saying 'I am hungry'. In short, Fregean thoughts are the entities that are said to be known and understood from utterances—the contents of I-utterances. The result that both the thinking animal and the Lockean person know the same Fregean thought so far follows from the neo-Lockean's commitment that they are

cognitively identical: what the Lockean person knows or understands is the same as what the thinking animal knows or understands.

4.3 Fregean Analysis: *de se* Knowledge

On the other hand, Noonan does acknowledge that there is a difference between the animal's knowing *that I am so-and-so* and the Lockean person's knowing *that I am so-and-so*: only the Lockean person knows *de se that I am so-and-so*. Let the term '*de se* knowledge' be defined by such difference. In this subsection we shall see how *de se* knowledge is characterised in terms of the current Fregean analysis. On *de se* knowledge, Noonan remarks:

"(...) Since I understand this sentence I also know the proposition thus expressed, and in doing so I know, *of myself*, that I am a person. But no more than this can be required for knowledge *de se*, of oneself, that one is a person.

The animal, on the other hand, does not know *of itself* that it is a person, since it is not. It does, however, know the thought it would express by saying 'I am a person'. Does it know that it is itself a person? Is it correct to say 'the animal knows that it itself is a person'? Yes: 'it itself' is here the indirect reflexive. It is not an expression whose reference is the animal (...) So it is true that the animal thinks that it itself is a person. But it is not true that the animal thinks *of itself* that it is a person, ..." (Noonan 2010b p.198-9)

I interpret the last statement "it is true that the animal thinks that it itself is a person. But it is not true that the animal thinks *of itself* that it is a person" as follows: since the Lockean person and the thinking animal have the same Fregean thought $T_{\gamma=\text{ego}(\text{LP})}^{\text{HA}} = T_{\gamma=\text{ego}(\text{LP})}^{\text{LP}}$, it entails that there is a same sense of 'I' that constitutes such Fregean thought— γ . γ is a sense of 'I' that is a mode of presentation of the person—the Lockean person. Senses determine references, and the same sense cannot differ in its reference. The phrase 'thinking of' indicates whether the reference of x 's thought is about x itself. The sense of 'I' that constitutes the animal's Fregean thought is a mode of presentation of the person, the Lockean person on Noonan's view, and thus the animal is not thinking of itself. The condition for thinking *de se* is the following:

- **De Se Knowledge**

x knows *de se* that P (thinks of itself) iff the reference of P is x (or the sense of 'I' is a mode of presentation of x)

This captures Noonan's characterisation of *de se* knowledge. As for the question: why is the sense of 'I' γ that constitutes the animal's Fregean thought not a mode of presentation of the animal? In other words, what determines γ is a mode of presentation of the Lockean person but not the animal? This is where Noonan's revisionism joins in, which claims that the rule for the reference of 'I' is the naïve rule plus the *personhood requirement*: i) x is the thinker of the I-thought & ii) x is a person. This rule dictates the sense of 'I' γ must be such that it is a mode of presentation of *the person who thinks the I-thought*. For the sake of exposition, we may think of it as a *definite description*: 'the person who thinks the I-thought'²², and the reference of the thought is the object that satisfies such description. So,

(1) LP knows *that I am so-and-so*

(1') HA knows *that I am so-and-so*

(2) LP knows that *the person who thinks the I-thought* (sense γ) is so-and-so

(2') HA knows that *the person who thinks the I-thought* (sense γ) is so-and so

(3) LP knows that *LP* (reference of γ) is so-and-so

(3') HA knows that *LP* (reference of γ) is so-and-so

(4) LP knows *de se* that P

(4') HA knows that P

How can we account for such difference between LP's knowing and HA's knowing that P appearing in (4) & (4')? To maintain coherency, it seems that

²² I suppose that the context of the discourse is clear enough to determine which I-thought is "the I-thought" in question.

Noonan must insist that knowing *de se* that P or not would not make a difference in one's understanding of the proposition. If there is any difference, it must be somehow *external* to one's understanding of the I-utterance. Or the *de se* qualification of the content P must be an *extrinsic* property to LP's cognition²³. Alternatively, if we speak of the propositional attitude *knowing* as a relation between a knower and a proposition, then one can suggest that knowing *de se* is a property attributed to the relation, not to the proposition as content. What prompts suggestions along this line is the neo-Lockean claim that in the *too many thinkers case*, the non-person-thinker and the person-thinker are cognitively the same. To account for the difference, we must appeal to something other than entities that analysed as ones that are responsible for the cognition of I-utterances.

4.4 Some Problems Concerning the Fregean Analysis

In the Fregean analysis, the contents of I-thoughts are analysed in terms of Fregean thoughts. In doing so, we are supposing a theory of propositions: propositions are complexes of Fregean senses. Understanding contents in terms of Fregean propositions is one way in which we can cash out Noonan's *personal pronoun revisionism*. In this sub-section, I shall outline some weaknesses specific for this Fregean approach to the defence of neo-Lockeanism.

The first worry has to deal with the general epistemology that this Fregean analysis carries over. x's believing that P is necessary for x's knowing that P. So, if it can be said that x knows that P, x must believe that P. According to Frege, in belief context the objects that stand in the *believing* relation are thoughts such that (i) thoughts are complexes of senses; and (ii) thoughts are bearers of meaning and understanding i.e. are responsible for the cognition of language. In other words, we gain knowledge via the meaning of the words. This seems to echo Dummett's conception of Frege: Frege replaces epistemology by the philosophy of language as the first philosophy²⁴.

²³ For example, the property of *being a brother of* is an extrinsic property for me since the obtainment of this property is determined by individuals other than myself: In a world *w* in which I have a sister, I have such property; in a world *w** in which I am the only son of my family, I don't have such property, but the me in *w* and the me in *w'* are the same individual. An example of an intrinsic property is the property of *being fat*.

²⁴ Dummett 1981, p. 669

Admittedly, when it comes to how we can know something about non-existing objects, explaining our knowledge of objects of this sort in terms of senses seems to be compelling. After all, the objects don't exist (or, as we may suppose, the nature of their existence is radically different than that of some everyday material objects like tables and apples, which are references in the current Fregean analysis). But if persons are indeed certain objects that exist in the world and one of them is numerically identical with me, claiming that I can only know something about such object *via* some senses i.e. modes of presentation, the same way as I know something about non-existing objects like unicorns and the king of France doesn't seem to be as compelling as the case of non-existing objects²⁵. Intuitively, there seems to be a more direct way, perhaps a relation between a knower and an object determined by a Fregean sense. But epistemology of this sort seems to be ruled out by the Fregean analysis of propositions.

A more compelling alternative for the epistemology of persons seems to be one that is suggested by Russell according to which propositions are analysed in terms of objects, properties, and relations, that is, the *theory of acquaintance* (Russell 1912), (1910): when I have a thought *that I am hungry*, I have a *de re* belief of an object, myself, to which the 'I' in the I-thought refers and which I have direct acquaintance with. In doing so, senses are not necessary.

The second issue is concerned with Frege's thesis about *de se* attitudes. Frege famously held that the propositions identified by I-utterances i.e. *de se* propositions are *private*:

"Now everyone is presented to himself in a particular and primitive way, in which he is presented to no-one else. So, when Dr. Lauben thinks that he has been wounded, he will probably take as a basis this primitive way in which he is presented to himself. And only Dr. Lauben himself can grasp thoughts determined in this way."

(Frege 1956, 298)

²⁵ I think explanation of this line would gain more plausibility if one is ready to admit that persons are in an important sense non-existing objects, thought, it offends the common sense that we exist.

However, Noonan must reject that *de se* propositions are private. It is because the *too many thinkers case* implies the sharing of thoughts between the temporary coincident thinkers. The thoughts, including I-thoughts, are in this sense 'public', at least public relative to the thinkers who share the thoughts. To preserve the coherence of neo-Lockeanism, Noonan must not follow Frege but claim the opposite: *de se* propositions are not private thoughts (they are, at least, locally sharable among the thinkers of them).

My last worry is a follow-up of the second one. It has to do with the characterisation of *de se* knowledge under the Fregean analysis Noonan suggests. His proposal seems to suggest a certain form of *externalism* on *de se* knowledge. As I explicated previously, one's knowing *de se* that P is determined by something external to one's understanding of P. Whether x's knowing *de se* or not that P makes no internal difference to x's cognition.

The problem is that *de se* knowledge understood this way seems to fail to capture some of the most important characteristics we normally attribute to it. *Prima facie* there seems to be *internal* differences between one's knowing *de se* that P and one's merely knowing that P. That is to say, when x's thought is a thought about herself, x will behave differently and we can only explain this by the appeal to some internal differences between x's merely knowing and knowing *de se*.

For example, suppose both Anne and Bongo have the thought *that Bongo wants to eat ice-cream*, a thought that Bongo would express by the utterance 'I want to eat ice-cream' and Anne 'He wants to eat ice-cream'. Put in Noonan's terminology both Anne and Bongo are thinking *of* Bongo via the sense expressed by the proper name 'Bongo'. However, we would expect Bongo to be the one, not Anne, who tries to get some ice-cream. The difference between their knowing *that Bongo wants some ice-cream* explains their different behaviours. Bongo seems to be motivated by some element in his thought, which is missing in Anne's. If we ask Bongo to explain his behaviour, he would say "It is because *I* want to eat ice-cream", but Anne would say "I am not the one who wants to eat ice-cream. He is." It is hard to see how one can explain the behavioural differences between Bongo and Anne without making reference to their internal cognitive lives such as

different desires, mental states, beliefs etc., and these have to do with whether the subject in Bongo's thought is himself—whether he knows *de se* that P or not. This is a strong motivation for claiming that there are internal differences corresponding to whether one's thought is about herself. If this is so, in the coincident case we should expect that the animal's knowing *that so-and-so* and the Lockean person's knowing *de se that so-and-so* should have certain internal difference as well. But this last claim not only contradicts the neo-Lockean's commitment that the Lockean person and the thinking animal are cognitively identical but also is one that the resources provided by the Fregean analysis could not explain.

4.5 Return to the Epistemic Problem

I have considered issues that are specific for Noonan's further explication of *personal pronoun revisionism* in terms of Fregean propositions. But this Fregean conception of propositions is just one of the theories of proposition in which the defender of neo-Lockeanism can give a full account of his *personal pronoun revisionism* as a response to the *epistemic problem*, and there exist different theories of propositions anyway.

What is essential to the solution is that it allows us to *resolve* the *epistemic problem*: when I ponder the question “which one of the thinkers is *me*?” or “am I the thinking animal or the person?”, the question appears to be *genuine* only because we are supposing a rather naïve rule for the reference of ‘I’, that is, ‘I’ refers to the thinker of the thought. In the coincident case, such rule will pick out the person-thinker as well as the non-person-thinker, of which posing the *epistemic problem* makes sense. It appears, as the naïve rule suggests, to be possible that the ‘me’ / ‘I’ in my thought can refer to the coincident thinking animal as well as the Lockean person. But according to Noonan, this naïve rule is not *the* rule for the reference of ‘I’. Instead, the rule for the reference of ‘I’ is a more restrictive rule, with the *personhood requirement* as an extra constraint on the object of reference. If this is the case, then there is no epistemic worry that whether the ‘I’ / ‘me’ in my thought refers to the person-thinker or the non-person-thinker since the ‘me’ or ‘I’ in my thought must refer to the person-thinker, and on the neo-Lockean account only Lockean persons are persons. The thought *that I*

am a person you have is trivially true: the ‘I’ refers to the person and you are a person. The revised rule, if true, resolves the *epistemic problem*.

The substantial claim of this proposal is thus the revised rule, which comes with the *personhood requirement*. That this further restriction on the reference of ‘I’ is essential to resolving the *epistemic problem* is obvious: if we drop this extra *personhood requirement* and employ the naïve rule, then the thinking animal’s I-thoughts would be thoughts about the animal, and the Lockean person’s thoughts about the Lockean person, a result that obviously goes against Noonan’s claim that both the animal and the Lockean person are thinking about the person. With the revised rule, Noonan can retain the claim that they are cognitively identical and the claim that the thought *that I am a person* is trivially true²⁶.

In response, there are two options in general (i) accept the proposal but argue against the neo-Lockean claim that persons are psychological continuers; (ii) reject the proposal *per se*²⁷. I shall take the second option and develop a challenge to Noonan’s *personal pronoun revisionism*. But this is not to say that I am trying to defend animalism by arguing against neo-Lockeanism since my main target is coincident thinkers. It could be, however, but it is not my main interest here²⁸.

5. Against Personal Pronoun Revisionism

My general strategy is to employ literature on the problem of *de se* attitudes in order to argue against Noonan’s *personal pronoun revisionism*. Having said that my target is not the specific version that Noonan defends but a more general and neutral one: i) it is neutral to neo-Lockeanism and animalism²⁹; ii) it presupposes a

²⁶ This point is clearly put in Noonan (2001) “I am a person” is analytic (...) If, as the referential ambiguity approach requires, I have my “rivals”, sharing my skin, they will indeed think my thoughts. But the thoughts they express using our shared vocal cords in the form “I am . . .” will in no case be thoughts about themselves but rather thoughts about me.” The Fregean analysis is an extended defense of the same idea.

²⁷ Noonan’s personal pronoun revisionism but free of the neo-Lockean claim that all persons are psychological continuers

²⁸ It could be, if what the animalist opposes to neo-Lockeanism is the claim that there are coincident person- and non-person-thinkers temporarily occupying the same temporal-spatial location and being capable of contemplating exact same I-thoughts.

²⁹ Again, the *personhood requirement* by itself doesn’t entail a theory of personal identity. If it turns out that persons are psychological continuers, then only the Lockean person can be said to know

theory of propositions but doesn't commit to any particular view on the nature of propositions³⁰.

In the discussion of *personal pronoun revisionism*, the neo-Lockean theorists' claim that the coincident thinking animal and the Lockean person are cognitively identical, is of special importance, for it poses a general constraint on the *contents* of their thoughts. Hence, if we are to say that there is a difference between their I-thoughts, such a difference cannot be explained by the appeal to different contents in their thoughts. However this creates the tension I briefly covered by the Anne/Bongo example in Section 4.4.

Noonan himself is aware of the discomfort that goes with his explanation³¹. We normally think that the contents of your I-thoughts and those of mine must be somehow different: x's I-thoughts are ones about x, and y's I-thoughts about y. How can x's thought *that I am so-and-so* be a thought about someone else other than x itself? It seems that x in virtue of being a thinker of an I-thought is *sufficient* for x to be the reference of the I-thought, and thereby x thinks *de se that I am so-and-so*. However, according to *personal pronoun revisionism*, this intuitive correspondence between being a thinker of an I-thought and being able to think *de se* is not necessarily accurate. The thinking animal in the *too many thinkers case* is a counter-example according to the neo-Lockean theorist.

My argument against *personal pronoun revisionism* expands precisely upon this discomfort. The worry is not (just) that it goes against our intuition regarding the correspondence between thinkers of I-thoughts and being able to think *de se*, but Noonan's acknowledgement that the difference between thinking *de se* that P and thinking that P is *not* a difference in content. Intuitively, such *de se* qualification *should* be one that is *of* the content, not merely one *of* the relation between the knower and the content that comes *extrinsically* to one's grasp of the

de se that I am so-and-so since the Lockean person is the reference of 'I'. The same goes for theories of the biological approach: if persons are biological continuers, then only the thinking animal can be said to know *de se that I am so-and-so* since the thinking animal is the reference of 'I'. It all comes down to the question what persons, fundamentally speaking, are.

³⁰ Propositions can thus be Fregean propositions, Russellian propositions or possible worlds

³¹ "Discomfort may remain. According to my version of the neo-Lockean account, it is not possible for something that is not a psychological continuer to be an object of its own first-person thoughts, no matter how sophisticated its thoughts are" (Noonan 2010c, p.97) The same goes for the animalist who accepts Noonan's proposal but rejects the neo-Lockean claim that persons are psychological continuers.

content (as it is suggested in Noonan's Fregean characterisation). As for the question why Noonan could not embrace the claim that the person-thinker and non-person-thinker have *de se* thoughts, the answer is straightforward: i) that the coincident thinking animal and the Lockean person are cognitively identical, must be accommodated and ii) the same Fregean sense cannot differ in its reference (but not vice versa).

In this section, I will attempt to argue that i) both thinkers are capable of having *de se* I-thoughts, and ii) the contents of their *de se* I-thoughts are different. Ultimately, I claim that we have enough motivation for rejecting Noonan's neo-Lockeanism.

5.1 Setting the Scene

In this sub-section, I shall outline and fix the meanings of some relevant terms, adopting the labels which will provide easy reference. I shall call the rule advocated by Noonan as 'revised rule for first-person reference' (RR) and the more intuitive rule as 'naïve rule for first-person reference' (NR).

- **Revised rule for first-person reference (RR)**

x is the reference of 'I' iff (i) x is a person; (ii) x thinks an I-thought identified by the I-utterance 'I am so-and-so'

- **Naïve rule for first-person reference (NR)**

x is the reference of 'I' iff x is a thinker of an I-thought identified by the I-utterance 'I am so-and-so'

I will speak of *propositions* and *contents* interchangeably. In general, I take propositions as the bearers of truth-values and cognitive significance of utterances. They are the entities that are expressed by the utterances. The term 'content' in my use is thus an umbrella-term encompassing Fregean propositions, Russellian propositions, possible worlds etc.. It is the general theory of propositions with which my main argument is concerned.

Since the sort of propositions that will be occupying our attention are the contents of our I-thoughts that correspond to I-utterances in particular, let us label them as 'indexical beliefs'.

- **Indexical beliefs**

Indexical beliefs are the kind of contents that correspond to indexicalised utterances i.e. I-utterances, which take the form 'I am so-and-so'.

I-utterances are characterised as utterances that contain the first-person pronoun 'I' and express propositions. At this moment I remain neutral about whether indexical beliefs are *essentially* indexicalised. This will be discussed later.

We can hold different propositional attitudes towards the proposition expressed by an utterance, and each one of them is characterised according to the word we used in making such belief reports. For example, if the belief report is 'I know that Anne has had ice-cream today', the propositional attitude is characterised by the word 'know' in my report, and is a relation between me and the proposition expressed in the that-clause in the attitude word i.e. the proposition identified by the that-clause 'that Anne has had ice-cream today'. Such relation is thus a *knowing* relation. There are other relations are, for instance, *believing*, *thinking*, *desiring* etc.

- **Knowing as a propositional attitude**

A holds the propositional attitude *knowing* towards her indexical belief iff (i) there is a *knowing* relation held between an agent x and a proposition P; (ii) $x=A$; (iii) P = an indexical belief identified by an I-utterance

In the coincident case, the animal and the Lockean person are said to be cognitively identical. This entails that the thinking animal and the Lockean person would (i) have the same indexical belief, and (ii) hold the same propositional attitude towards that belief.

- **Cognitively identical (CI)**

The Lockean person and the coinciding thinking animal are cognitively identical iff they can grasp the same contents when understanding I-utterances, and hold the same propositional attitude towards them.

Since CI will play an important role in my argument, I shall explain a bit more about the notion of CI I will be discussing. It is a *materialist* notion. When Noonan tries to explain why the animal cannot refer to itself like the Lockean person, he makes the point with reference to materialism.

“To the question ‘Why is the animal not thinking of itself in the first-person way given that the coincident person is?’, (...) from the materialist viewpoint shared by the animalist and the neo-Lockean it is incomprehensible that things like the coincident person and animal, which are microphysically indistinguishable throughout the relevant period (differing only in irrelevant far off past or future respects), could be cognitively distinct, could be thinking about different things or thinking of them in different ways; the mystery would be if they did.”

(Noonan 2012, p. 318)

The appeal to identical material constitution explains why the Lockean person and the thinking animal must have the same contents in their thoughts throughout the relevant coinciding period. This will be the notion of CI I take, which is agreed by both Olson and Noonan.

Lastly, we have the notion of *de se* attitude.

- **De Se attitudes (DSA)**

x holds a *de se* attitude A towards a proposition P iff (i) there is a relation characterised by attitude A held between x and a proposition P, (ii) P is identified by an I-utterance, and (iii) P is about x itself.

Let us focus on (iii) for the moment. Depending on one’s view on propositions, the statement “P is about x” can have quite different interpretations. According to the Fregean view we have seen in Section 5, a proposition is a complexes of Fregean senses i.e. a thought. A sense is a mode of presentation of

an object that can be understood as a set of descriptive conditions. So the referent, as an object of which the sense of 'I' is a mode of presentation, is the entity in (iii) that determines whether P is about x. On the Russellian view, by contrast, a proposition is constituted by objects and properties (and relations). "P is about x" means that there is a relation between x and proposition P such that the subject that constitutes proposition P is x (not merely a mode of presentation of x). DSA as such is by no means a definition of *de se* attitudes but a first approximation.

5.2 The Argument

The crucial element in my argument is related to the general discussion of the problem arises from *de se* attitudes, originally called "the problem of the essential indexical" in Perry (1979).

The gist of my main argument is the following: mainly due to his adherence to CI³², Noonan's proposal also suffers from the problem of *de se* attitudes in general discussed in the literature. In addition, according to the received view largely motivated by that problem, both the animal's thinking that P and the Lockean person's thinking that P are, contrary to Noonan's account, thinking *de se*. It follows that their indexical beliefs are also different in content, for they are about different things. Hence, they are not cognitively identical thinkers, which leads to the refutation of neo-Lockeanism on Noonan's own terms.

The structure of my main argument is as follows:

- **The argument against the revised rule of the reference of 'I' (RR)**
 1. If RR is the rule for the reference of 'I', then only one of the thinkers in the coincident case who satisfies the *personhood requirement* by RR is thinking *de se* that P
 2. According to the received view on *de se* attitudes, RR is not the rule for the reference of 'I', but NR is.
 3. According to NR, both thinkers are thinking *de se* that P.

³² I take Noonan's Fregean analysis and the related characterisation of *de se* thoughts are mainly prompted by CI.

RR is the key to resolving the *epistemic problem*. If it turns out, as the argument shows, that we have independent reason for favouring NR over RR, the problem returns: for all I know, I am able to think *de se that I am a person*. However according to the received view, the thinking animal and the Lockean person are both able to do so. How do I know which one of them is me? (In comparison, there would be only one of them who is able to think *de se* according to RR). By NR, 'I' can refer to the person-thinker as well as the non-person-thinker. For all I know, I can be either one of them, and my utterance 'I am a person' is true if I am the person-thinker; false otherwise. Given CI and the claim that both of them are capable of thinking *de se* how can I know which one of them is me? How can I individuate one of them as the referent of my I-utterance? The first argument, then, if successful, leads to the return of the *epistemic problem*.

Building on top of the first argument, here is the second argument.

- **The argument against CI**

1. Suppose CI is true
2. It follows from CI that the indexical beliefs that the Lockean person and the thinking animal have must be the same in content
3. According to the received view, both of them are thinking *de se* that P
4. Given (3) and the notion of *de se* content the indexical beliefs that the Lockean person has must be different than the indexical beliefs that the thinking animal has in content
5. (4) contradicts (2)
6. Therefore, CI is false

Premise (3) is just the conclusion of the first argument. What is new is the notion of '*de se* content' (a notion substantiated by the received view on *de se* attitudes) premise (4) appeals to. The conclusion of the second argument is just one step away from the refutation of Noonan's neo-Lockeanism: if the contents of their *de se* thoughts are indeed different, then how can it be said that they are cognitively identical? If one is to give up CI, the acceptance of the *too many thinkers case* is not motivated, at least from a materialist point of view: how can

two temporally coincident thinkers composed of exact same material matters differ in their thoughts? One possible way to defend the position is to bite the bullet: one may endorse that they are indeed thinking about different things, but argues for another sense of CI:

- **Cognitively parallel (CP)**

The thinking animal has its own stream of thoughts, and so does the Lockean person while having indexical beliefs. But, these two streams of thoughts, at least in the case of indexical beliefs, are *private* to their owners in the sense that only the thinker who thinks *de se* that P can entertain the indexical belief.

CP has its roots in Frege's view on private thoughts. The reason for introducing CP may be to allow the neo-Lockean theorist to maintain something like CI and the *too many thinkers case* coherently. The notion of private accessibility of one's indexical beliefs doesn't seem to be obviously absurd, if not defensible, and they can still share non-first-person thoughts. CP allows us to resolve the *epistemic problem* in the following way: The thought *that I am a person* I have is *de se* and true. In the coincident case, the person-thinker's private thought *that I am a person* is *de se* and true; the non-person thinker's private thought *that I am a person* is also *de se* but not true since the non-person thinker is not a person. We can individuate non-person-thinker from person-thinker according to the truth and falsity of the propositions. But if one is to maintain CI, it is still difficult to see how one can *know* that she is the one thinking truly that she is a person. After all, by SI, the truth and falsity of the contents will give one the same motivation for behaviour and understanding of the contents.

There is a more pressing issue with CP, that is, how can we make sense of CP? The appeal to *dualist* ontology may do the trick. However, this is not an option for Noonan who endorses a *materialist* ontology³³, and I wonder how much this can help the neo-Lockean theorist in the face of *the epistemic problem*. Therefore,

³³ "...from the materialist viewpoint shared by the animalist and the neo-Lockean it is incomprehensible that things like the coincident person and animal, which are microphysically indistinguishable throughout the relevant period (differing only in irrelevant far off past or future respects), could be cognitively distinct, could be thinking about different things or thinking of them in different ways; the mystery would be if they did." (Noonan 2012, p. 318)

I conclude that if the falsehood of CI is established, and we have a strong motivation for abandoning the materialist version of neo-Lockeanism (if not the dualist version altogether).

I now turn to the presentation of the received view on *de se* attitudes on which my arguments depend. Since the received view is characterised by problems, we shall take a little detour before we get to the view on *de se* attitudes.

5.3 The Received View on *De Se* Attitudes

The so-called received view on *de se* attitudes is one that corresponds to certain problems risen from *de se* attitudes. So regarding a received view, there is also a 'received' problem on *de se* attitudes. Perry (1979) presents such a problem, calling it 'the problem of the essential indexical'. Ninan (2015) notes: "the idea that *de se* attitudes pose a challenge to theories of attitudes is now the received view". A little later she adds: "the literature on these topics has been less than completely clear on just what that problem is supposed to be" (p.1). Getting clear what the problem of *de se* attitudes really is is not my task in this thesis. For the sake of my argumentation, I am concerned with the claim that Noonan's account leads to the same problem that the received view on *de se* attitudes is after.

I will outline two 'received' problems on *de se* attitudes, calling them respectively '*the problem of the essential indexical*' and '*the problem of de se attitudes*'³⁴. The former leads to the claim that indexicals are essential, supporting premise (2) of the first argument; the latter leads to a substantial account of *de se* content, which is a notion employed in premise (4) of the second argument.

We shall begin with the doctrine of propositions since it is the target of the two problems. To better capture the point, I employ the expanded version of Perry's original presentation of the doctrine (1979, p.5-6) by Ninan (2015)³⁵:

- **Three tenants of the doctrine of propositions**

³⁴ I call them as two problems because they are illustrated by two different examples: the *messy shopper case* and the *bear attack case*. I think a full account on *de se* attitudes will show that these are just but one problem.

³⁵ Another advantage of employing Ninan's version is that it fits better with my use of the word 'content'. Perry uses the word 'belief' instead e.g. "The first is that belief is a relation between a subject and an object, the latter being denoted, in a canonical belief report, by a that-clause." (p.6)

“(1) Two-Place Relation

An attitude relation (believing, desiring, etc.) is a two-place relation between an agent and a content.

An attitude (type) like the belief that snow is white is a mental state that consists in having a certain relational property, i.e. the property of standing in the relation of believing to the content that snow is white.

(2) Frege's Constraint

Contents are assigned to attitudes in a manner that accommodates Frege cases. So if a rational agent could have a belief he could express by saying, “Hesperus is bright” without having a belief he could express by saying, “Phosphorus is bright,” these two beliefs have different contents.

(3) Absoluteness

The contents of attitudes are absolute, i.e. contents do not vary in truth value across individuals or times.”

(Ninan, 2015, p.7)

● Two expansions

(4) Agreement

Agreement is a two-place relation between a group of individuals and a content.

A group of individuals stands in the agreement relation to a content p iff all the members of the group believe p .

Let B_{xp} and B_{yq} be token beliefs belonging to x and y respectively. Then x and y agree on something in virtue of x 's having token belief B_{xp} and y 's having token belief B_{yq} iff $p = q$.

Let D_{xp} and D_{yq} be token desires belonging to x and y respectively. Then x and y agree on how they'd like things to be (in one respect) in virtue of x 's having token desire D_{xp} and y 's have token desire D_{yq} iff $p = q$.

(5) Explanation

Suppose the fact that x performed action α is explained by the fact that x has beliefs B_{xp1}, \dots, B_{xpn} and desires D_{xq1}, \dots, D_{xqk} .

Then, if y has beliefs B_{yp1}, \dots, B_{ypn} and desires D_{yq1}, \dots, D_{yqk} , then, other things being equal, y will also perform α .

(Ninan 2015, p. 15-7)

Our first problem, the *problem of the essential indexical*, has to do with tenant (3) ABSOLUTENESS and EXPLANATION, illustrated in Perry's *messy shopper case*:

- **Messy shopper case**

"I once followed a trail of sugar on a supermarket floor, pushing my cart down the aisle on one side of a tall counter and back the aisle on the other, seeking the shopper with the torn sack to tell him he was making a mess. With each trip around the counter, the trail became thicker. But I seemed unable to catch up. Finally it dawned on me. I was the shopper I was trying to catch."

(Perry 1979, p.1)

In such a case, we may assume that there is a belief, *B*, that Perry came to believe and led him change his behaviour from looking down at the floor to straightening up and checking his own cart. *B* is the crucial belief that leads Perry to change his behaviour and is identified by his utterance 'I am making a mess', that is, the belief *that I am making a mess*. However, the belief *that I am making a mess* identified by the utterance 'I am making a mess' is not true or false absolutely: "had another shopper said it when I did, he would have been wrong. So the sentence by which I identify what I came to believe doesn't identify, by itself, a proposition." (p.6-7). The proponent of the doctrine of propositions will need to replace 'I' with something that fits Perry the shopper alone and, in accordance with tenant (3), the semantic value of the resultant sentence by such replacement should not vary relative to different speakers. The replacement of 'I' is thus "a missing *conceptual ingredient*: a sense for which I am the reference, or a complex of properties I alone have, or a singular term that refers to no one but me" (p.7).

- **Conceptual ingredient**

Replacement: 'I' / 'α'

Utterance: 'I am making a mess' <-> 'α is making a mess'

Belief report: *that I am making a mess* <-> *that α is making a mess*

Replacing the indexical 'I' with α as a missing conceptual ingredient the resultant utterance ' α is making a mess' should identify a proposition that Perry in the *messy shopper case* believes. Let us consider two obvious candidates for α : the proper name 'John Perry' and the definite description 'the bearded philosopher in a Safeway store west of the Mississippi'.

- **Proper name**

Utterance: 'John Perry is making a mess'

Belief report: I (Perry) believe *that John Perry is making a mess*

- **Definite description**

Utterance: 'The bearded philosopher in a Safeway store west of the Mississippi is making a mess'

Belief report: I (Perry) believe *that the bearded philosopher in a Safeway store west of the Mississippi is making a mess*

The problem is that the beliefs, the propositions, rendered by such replacements, don't seem to identify the crucial belief B that leads Perry to change his behaviour. For they differ in their explanatory power for the causes of behaviours. For example, Perry could believe that *that John Perry is making a mess* without changing his behaviour because he might forget his own name. He would, however, if we supplement the following belief *that I am John Perry*, but it brings in the indexical again. The same goes in the case of definite description. By EXPLANATION, the resultant beliefs rendered by the above replacements don't have the same content that B is supposed to have, unless we are willing to supplement a further belief *that I am α* . Therefore, we have a case that the indexical element in B is *essential* in the sense that it is irreplaceable in our belief reports, a view on *de se* attitudes that corresponds to the *problem of the essential indexical*.

The second problem, *the problem of de se attitudes*, has to deal with AGREEMENT and EXPLANATION, which is illustrated in the *bear attack case* by Perry (1977):

- **Bear attack case**

“...imagine that you and I are walking in the woods when a bear begins to chase me. Now suppose that you and I agree on all of the relevant facts of the case, and we also agree on how we'd like the situation to turn out. For example, suppose that we agree that I'm being chased a bear. That is, suppose I believe *de se* that I'm being chased by a bear and that you believe *de te* that I am being chased by a bear, i.e. you have a belief you could express to me by saying, “You are being chased by a bear.” Now if, as I am assuming, we count as agreeing that I am being chased by a bear in virtue of my having this *de se* belief and your having the corresponding *de te* belief, then it would seem to follow from AGREEMENT that these beliefs have the same content. So if we let p_1 be the content of my *de se* belief, p_1 is also the content of your *de te* belief. Letting ‘m’ denote me and ‘y’ you, this means that I have a token belief $B_{p_1}^m$ and you have a token belief $B_{p_1}^y$. ”

(Ninan 2015, p.16-7)³⁶

The problem is that AGREEMENT and EXPLANATION yield incompatible results. The example in the quote shows that by AGREEMENT my token belief $B_{p_1}^m$ should have the same content as your token belief $B_{p_1}^y$. Putting it in another way, you agree with me on the sentence ‘I am being chased by a bear’ uttered by *me*. The content identified by ‘I am being chased by a bear’ is the belief *that I am being chased by a bear*. In virtue of agreeing with me, you believe that *that I am being chased by a bear*, a belief you would normally express by the utterance ‘you are being chased by a bear’. Now, suppose further that the belief *that I am being chased by a bear* leads to the behaviour *curling into a ball*. By EXPLANATION both you and I should perform the same behaviour *curling into a ball* since we have, by agreeing with my utterance “I am being chased by a bear”, the belief that *that I am being chased by a bear*. But this is wrong. I am the only one who is being chased by the bear and should be the one who curls into a ball to survive the

³⁶ The case is originally from Perry (1977, p.23)

attack. You as a clear-minded and good-hearted person should run away and call for help.

One way to look at it is that 'I am being chased by a bear' uttered by me and 'You are being chased by a bear' uttered by you don't identify the same *de se* belief *that I am being chased by a bear* we thought they did initially. You agreeing with my utterance doesn't follow that you and I have the same contents in our beliefs. This line of thinking suggests us to retain EXPLANATION instead of AGREEMENT. The proposition I come to believe is *that I am being chased by a bear*—a *de se* belief; a different proposition you come to believe is *that you are being chased by a bear*—a *de te* belief (compare previously we said that your belief is also the belief that *that I am being chased by a bear*—an apparently *de se* belief. The problem of *de se* attitudes illustrated by the *bear attack case* brings us to a more substantial notion of *de se content*.

- **De Se Content**

“(1) If x's thinking *de se* that *p* would lead to the behaviour *b*, and y's thinking *de se* that *q* would not lead to the behaviour *b*, then the token belief *p* that x has is different than the token belief *q* that y has in content

(2) If x's token *de se* belief *p* is such that it would lead to behaviour *b*, and the token belief *q* that y has is the same in content with *p* ($p = q$), then y's token belief *q* is a *de se* belief that would lead to behaviour *b*.”

Ninan (2015, p.27-8)

So here we have two problems concerning *de se* attitudes. They are problems for the expanded doctrine of propositions as such. Now I shall argue that they are *also* problems for Noonan's *personal pronoun revisionism* by showing that he is committed to the relevant tenants that give rise to the two problems, the *problem of the essential indexical* and the *problem of de se attitudes*.

The *problem of the essential indexical* has to do with ABSOLUTENESS and EXPLANATION. We said that the utterance 'I am making a mess' doesn't identify a proposition because it is not true or false absolutely. By ABSOLUTENESS the proposition identified by the I-utterance must not vary its truth-value relative to different utterers. The resultant sentence 'John Perry is making a mess' with the 'I'

replaced by a proper name is a candidate because its truth-value doesn't vary relatively. Due to CI, Noonan cannot allow an I-utterance to differ its content relative to the coincident thinkers. In other words, 'I am so-and-so' by the thinking animal and 'I am so-and-so' by the Lockean person must express the same content. RR (understood as descriptive conditions) plays the role that proper names play in the example. The content expressed by the two coincident thinkers is identified by the sentence 'the person who thinks the I-thought is so-and-so'. Therefore, it is clear that Noonan is committed to a special version of ABSOLUTENESS that involves person- and non-person-thinkers.

As for the second problem, the *problem of de se attitudes*, which is concerned with AGREEMENT and EXPLANATION, we have something quite similar. Due to his acceptance of the *too many thinkers case* and CI, Noonan is committed to saying the Lockean person's and the thinking animal's thoughts are the same in content. His commitment to materialism rules out the possibility that the contents in the two coincident thinkers thinking could differ, which in turn grounds the claim that they are thinking the same set of token beliefs. This commits him to a more trivial version of AGREEMENT: the person-thinker and non-person-thinker agree on something in virtue of the person-thinker's belief is *numerically* identical with the non-person-thinker's (compare AGREEMENT defined by the phrase "...in virtue of x's having token belief B_{xp} and y's having token belief B_{yq} iff $p = q$ ", which suggests two *numerically* distinct but *qualitatively* identical token beliefs.)³⁷

The *problem of the essential indexical* and the *problem of de se attitudes* are also problems for Noonan's account. In what follows, I will take for granted the received view and show how it is in conflict with Noonan's *personal pronoun revisionism* characterised by RR and CI.

³⁷ Strictly speaking, one may consider that it is something other than AGREEMENT since we don't have two token beliefs agreed in the coincident case. But I think, with some reservation, that AGREEMENT is still applicable because what explains A agrees with B is such that A and B have the same belief, and I think in the coincident case it is correct to describe the non-person-thinker and the person-thinker "have the same belief". So what explains AGREEMENT will also explain this special version of AGREEMENT. It may be said that the phenomenon of agreement does not exist in the coincident case, but what explain this more authentic phenomenon of agreement should also explain the phenomenon, whatever it may be, of the coincident case.

5.4 Defending The Argument Against RR

As the *problem of the essential indexical* suggests, in the coincident case the I-utterance 'I am a person' by itself doesn't identify a proposition since it is not true or false absolutely: true if it is uttered by the person-thinker; false if uttered by the non-person-thinker. RR understood as a descriptive condition plays the role that the definite description that uniquely picks out Perry in the *messy shopper* case plays. With the indexical replaced, the resultant sentence 'The person who thinks the I-thought is a person' is now true or false *simpliciter* (trivially true in fact). Drawing on the result we achieved from the *messy shopper* case, such replacement would no longer preserve the explanatory force that the indexical belief has. By EXPLANATION, they don't identify the same proposition. We cannot replace 'I' with 'the person who thinks the I-thought' as suggested by RR.

On the other hand, NR understood as the following descriptive condition: the thinker of the I-thought can preserve the explanatory force that the indexical beliefs have. What I am suggesting is that in general 'I am so-and-so' and 'The thinker of the I-thought is so-and-so'³⁸ identify the same proposition, expressing the same content. Note that this is not the same as replacing by some definite description. The difference is that the definite description that can play the role of being a missing conceptual ingredient is supposed to be *absolute*, meaning that its semantic value will not change relative to the speaker or the thinker. But the descriptive condition by NR is different. The description 'the thinker of the I-thought' doesn't uniquely single out a reference. In the coincident case, it is not, for there is no unique thinker of the I-thought. Hence, given the essentiality of indexicals shown by the *messy shopper* case, NR is more preferable than RR as the rule of the reference of 'I'. If this is so, then in having the belief *that I am a so-and-so* both thinkers are qualified as thinking *de se* since both are thinking of themselves respectively. As the first argument goes, the *epistemic problem* returns.

5.5 Defending the Argument Against CI

With the conclusion of the first argument established, we can now claim, contrary to Noonan, that both the thinking animal and the Lockean person think *de*

³⁸ Or something alike, for instance, 'The speaker of...', 'The agent of ...', 'The believer of ...' etc..

se that P when having an I-thought viz. their thoughts are about themselves respectively. What we need to show in order to defend the second argument is that the Lockean person's *de se* belief and the animal's *de se* belief have different contents. Drawing on *De Se Content* followed from the *bear attack case*, I shall now attempt to show this.

The second argument begins with the premise that CI is true. As we saw from the previous discussion, this is the general constraint on Noonan's account that leads to the development of *personal pronoun revisionism*, characterized by the revised rule RR and the claim that only the person-thinker, the Lockean person on neo-Lockeanism, is capable of thinking *de se* when having an I-thought³⁹. In short, the coincident thinkers are both thinking truly about the person-thinker. But this result is in tension with the received view on *de se* attitudes followed from the *bear attack case*.

Consider the I-utterance 'I am hungry' and the corresponding belief *that I am hungry*, and suppose that the behaviour *buy ice-cream* will be caused by such belief. When the person-thinker says "I am hungry", both the non-person-thinker and the person-thinker have the belief identified by the utterance—the belief *that I am hungry*.

The person-thinker's *de se* belief *that I am hungry* will lead to the behaviour *buy ice-cream*. What about the non-person-thinker? By CI, we should expect that the non-person-thinker would have the same content in its beliefs, a situation that is similar to the *bear attack case* in which you and I both agree that I am being attacked by a bear. If the person-thinker's *de se* belief can motivate the behaviour *buy ice-cream*, we should expect the same for the non-person-thinker. Then, according to the second principle of *De Se Content*, the non-person-thinker's belief is also a *de se* belief. However, this last claim is in contradiction with Noonan's claim that *only* the person-thinker, the Lockean person, is capable of having *de se* beliefs. By reductio, as the second argument goes, we have a false assumption. RR and the claim that only the person-thinker is capable of thinking *de se* follow from CI. Therefore, CI is the false assumption.

³⁹ Noonan 2010b, p.198

5.6 Possible Objections

In this subsection I wish to consider three possible objections to my main argument.

Objection: The discussion of personal identity need not be framed with personal pronouns. The arguments from *de se* attitudes are not applicable to the discussion or only have a limited application to one particular framing of the discussion.

Reply: I agree that there exist indexical as well as non-indexical ways of framing the discussion of personal identity. So, we could side step the arguments from *de se* attitudes.

However, this is not an objection on Noonan's behalf, who thinks that i) the term 'person' *means* the same as personal pronouns such as 'I' and 'you'⁴⁰ ii) the formulation with personal pronouns is *primitive*⁴¹. Given (i) the meaning of personal pronouns must reconcile the meaning of the term 'person'; given (ii) the meaning of personal pronouns is more primitive. For those who endorse (i) and (ii), considerations from *de se* attitudes apply.

Objection: The application of the problems of *de se* attitudes to the coincident case is not appropriate. It is because there is a fundamental difference between the coincident case and the cases considered as ones that give rise to the problem of *de se* attitudes that is being dismissed: not all of the thinkers but only one of them is a person in the coincident case. As for the cases that reveal the problems of *de se* attitudes e.g. the *messy shopper case*, the *bear attack case*, these are cases such that the thinkers involved are presumed to be persons. The view that indexical beliefs are essentially indexical, which suggests NR as opposed to RR, seems true insofar as we are considering cases of the latter sort. The scope of the received view on *de se* attitudes only concerns persons. Indeed, in the coincident case the non-person thinker's indexical beliefs must, after all, about *something*. RR is introduced especially for determining the reference of the

⁴⁰ Noonan 2010c, p.96

⁴¹ "The debate to which the neo-Lockean intends to contribute is about two questions: 'What am I (fundamentally)?' and 'What are my persistence conditions?' The primary formulation of these questions is first-personal." (p.96)

indexical beliefs of non-person-thinkers. RR and NR apply to different sort of cases.

Reply: This objection fails. The difference, no doubt, is present, but it is irrelevant, unless one can show that being a person or not is relevant to a thinker's *cognition* of I-utterances. For example, it may be argued that persons must be able to think in a *more* sophisticated way than non-persons. If only the Lockean person is a person, it follows that the thinking animal must be a less competent thinker. NR applies only to thinkers whose capacity in this regard is on a par with persons'.

However, this line of defending is blocked by the neo-Lockean's commitment to CI: if the animal lacks certain capacity e.g. the capacity of having *de se* thoughts, so does the Lockean person, and *vice versa*; if the Lockean person has this capacity, so does the animal, and *vice versa*. The fact that one of the thinkers is a person and the non-person-thinker is cognitively identical with the person-thinker during the coincident period grants the applicability of the considerations for *de se* attitudes to the coincident case.

Objection: Various authors have written about the special problem concerning *de se* attitudes. Perhaps there exists some way in which we can accommodate the problems of *de se* attitudes without conflicting Noonan's account.

Reply: I consider the suggestions made by Perry (1979), namely i) relativised propositions, ii) limited accessibility, and iii) '*de se* belief states'. Of the three options, (iii) seems to be the most promising. Nevertheless, I shall show that mainly due to the neo-Lockean commitment to CI, none of them can be taken to retain neo-Lockeanism.

Relativised Propositions: The idea of the first option is that propositions are true or false relative to certain indexes. Hence regarding the belief *that I am making a mess*, instead of being true or false *simplicitor*, it is true or false relative to an index. In doing so, we give up ABSOLUTENESS. Applying it to the *messy shopper case*, we say Perry's belief *that I am making a mess* is true for an index—a particular person, Perry, the one who is making a mess. But, as noted by Perry, this doesn't seem to identify the crucial belief that led Perry to change his behaviour. Perry could believe that the belief *that I am making a mess* is true iff it

is true *relative to* Perry⁴² without changing his behaviour. For he might forget his name and thus wouldn't realise he himself is the index that he is calling "John Perry". In addition, "You believe so now, and doubtless have no more desire to mess up supermarkets than I did. But you are not bending over to straighten up a sack of sugar." (p.13). Perry would change his behaviour if the belief he has is such that *that I am making a mess is true for me*, which brings up the indexical again. By EXPLANATION, the non-indexicalised truth condition doesn't identify the same belief as the one with indexicalised truth condition.

Back to *personal pronoun revisionism*, suppose the truth conditions for propositions are relative to indexes. Originally, we say the animal's belief *that I am so-and-so* is true for *me*. By *personal pronoun revisionism*, the 'me' in the truth condition refers to *the person who thinks the I-thought* (RR) instead of *the thinker of the I-thought* (NR). But in doing so we will no longer preserve the explanatory force of the original indexicalised truth condition as we have seen before. This solution thus fails for independent reasons discussed. In fact we can modify the first argument from 'I' as content of an indexical belief to 'I' as an index of the truth condition for an indexical belief, and re-run the argument. This way, the solution would fail for related reasons in this modified argument⁴³.

Limited Accessibility: The second solution appeals to the limited accessibility of indexical beliefs. Given the constraint posed by CI and Noonan's commitment to materialism, it is hard to see how the temporally coincident Lockean person and thinking animal composed of exact same material substances

⁴² or to a definite description that picks out Perry alone; or he could have a *de re* belief *that someone is making a mess* in the his *mirror case*: "Suppose there were mirrors at either end of the counter so that as I pushed my cart down the aisle in pursuit I saw myself in the mirror. I take what I see to be the reflection of the messy shopper going up the aisle on the other side, not realizing that what I am really seeing is a reflection of a reflection of myself." (1979, p.12)

⁴³ Even if we take a step back and apply it to the coincident case anyway, it still fails. Consider the following relativised belief *that I am hungry* the truth condition of which is such that it is true for me. By RR, the belief is true for the person who thinks the I-thought. By CI, the animal also has the same belief *that I am hungry*, but it is true relative to the Lockean person and false relative to the thinking animal on the neo-Lockean view. We may expect the belief *that I am hungry* will lead to certain behaviour, for instance, *buy ice-cream*. If we take the perspective of the person, it seems right to expect her to perform *buy ice-cream*. The fact that her belief is true seems to explain her behaviour. Then, by CI, we should expect the non-person thinker has just the same motive for performing such action. But if we subscribe to the relativised proposition view, it seems to follow that it is her *false* belief *that I am hungry* that leads her to perform *buy ice-cream*. The truth and falsity of a belief cannot explain the behaviour caused, for the true belief and the false belief are the same in motivating one's behaviour. What else, given CI, can we appeal to then?

but nevertheless each of them has private indexical beliefs (this point was covered at the end of Section 6.2, I shall not repeat here).

Belief States: The idea of Perry's third solution is to reject AGREEMENT and introduce a new entity that can play the role that contents are supposed to play. Such entity is *belief states*:

"I propose we look at things in this way. The shoppers, for example, are all in a certain belief state, a state which, given normal desires and other belief states they can be expected to be in, will lead each of them to examine his cart." (p.18)

Perry, the shopper in the *messy shopper case*, is one of the shoppers in the quote whose behaviour was caused by such belief state. Since beliefs states play the role that we expect contents to play, they are, like beliefs (or propositions), content-like entities. So we have two content-like entities, beliefs and belief states⁴⁴. To see how AGREEMENT is rejected, we shall refer to the *bear attack case* once again. Both you and I agree that I am being traced by a bear. There is some content expressed by the utterance 'I am being traced by a bear' by *me* and the utterance 'You are being traced by a bear' uttered by *you* in which you and I mutually agree. This agreed content is captured by the content-like entity *beliefs*. By EXPLANATION, there must be some content difference in mine and your believing that explains our different behaviours. This is captured by the content-like entity *belief states*. So taking into account the totality of content-like entities mine believing and your believing don't have the same content, for they differ in belief states. Hence, AGREEMENT is rejected⁴⁵.

However, given the neo-Lockean commitment to CI, it is difficult to see how the temporary coincident thinkers can differ in their belief states. I conclude that despite the ingenuity of Perry's proposal, it is not one that the neo-Lockean theorist can employ.

⁴⁴ The underlying assumption is that believing is not a two-place relation between a believer and a content.

⁴⁵ Further elaboration, see Ninan 2015 section 4.2 p. 29-30

6. Conclusion

I shall sum up what I have shown and argued in this thesis.

In an attempt to refute the *thinking animal argument* as an argument against neo-Lockeanism which he favours, Noonan accepts the coincident case (*too many thinkers case*): there exist two thinkers, the Lockean person and the thinking animal, which are i) temporarily coincident entities and ii) cognitively identical (CI). According to the animalist Olsen, this will lead to the *epistemic problem*. Drawing on the distinction between thinkers of I-thoughts and objects of first-person reference, and the revised rule for the reference of 'I' RR, Noonan argues that the *epistemic problem* can be resolved.

I have argued in Section 5 that RR and CI are in tension with the received view on *de se* attitudes. For considering problems of *de se* attitudes, I claim that i) both coincident thinkers are capable of thinking *de se*, and ii) their *de se* beliefs have different contents. It follows from (i) that the *epistemic problem* returns, and from (ii) that CI is false. In the face of the *epistemic problem* and the falsehood of CI, I finally conclude that we have strong motivation to regard the neo-Lockeanism Noonan favours as not true.

Appendix: Noonan's Psychological Continuity Account

The Psychological Continuity View (neo-Lockeanism): Noonan's

position is that P1 at t1 is the same person as P2 at t2 if and only if P2 is psychologically continuous with P1 at t1. The psychological connections between P2 and P1 are "any causal links between past factors and present psychological traits", and "hold when a belief or a desire or any other psychological feature, persists" (1989/2003 p.10). Compare, the Memory Criterion by Locke, which states that P1 at t1 is the same person as P2 at t2 iff P2 has the memory of P1⁴⁶. Noonan's view inherits the psychological aspect of Locke's view, but in a more inclusive way, with emphasis on continuity.

Quinean Kind-membership Framework: Noonan suggests that the philosophical questions of a person's identity over time such as "what are the necessary and sufficient conditions for P1 at t1 to be the same person as P2 at t2?" should be rephrased to ones about an entity's *diachronic personhood*⁴⁷ such as "what are the necessary and sufficient conditions for an entity x to be a person from t1 to t2?"⁴⁸. More precisely, what we set to find out is the necessary condition for something to be a member of the kind we call 'Person'⁴⁹. According to Noonan, the necessary conditions for x to be a person is captured by the following de dicto principles:

'Passing-away' condition: necessarily, if x is a person then if x exists at t and
t* then Rxtt*;

⁴⁶ 'Consciousness' is the word Locke used in his original formulation. I interpret what he meant was memories.

⁴⁷ "This, then, I suggest, is what the request for 'a criterion of diachronic identity for K's', or an account of what constitutes K-identity over time, comes down to when properly expressed. It would perhaps be better described as a request for 'the diachronic criterion of K-hood'. 1989/2003 p.89

⁴⁸ It is because he follows Lewis (1986: 192-3) in thinking that the only identity relation is absolute identity (1989/2003) p.85, the specification of the satisfaction conditions of which is unproblematic, trivial and uninformative in that genuine problems cannot even arise. By rephrasing the question into an enquiry about diachronic personhood genuine questions such as how to informatively specify the criterion of personal 'identity' over time can arise.

⁴⁹ The original idea of which is taken from Quine (1976) but free of the commitment to the existence of temporal parts of physical objects (1989/2003 p.87)

Preservation condition: necessarily, if x is a person then if Rx_{tt^*} then x exists at t and t*

(Noonan 2010, p.84, schemas in Noonan 1989/2003, p. 88)

The phrase 'persistence conditions for persons' is construed as the specification of the relations R in the principles. Since this is what is at issue in the kind-membership framework, phrases such as 'criterion of personal identity over time', 'necessary and sufficient condition for P1 at t1 to be the same person as P2 at t2' shall be understood with reference to the above de dicto principles governing the kind called person.

Conceptual Enquiry: Noonan thinks that finding out the persistence conditions of persons is a matter of conceptual enquiry. Persons are therefore members of the extension of the concept *person*. To give an adequate analysis of the concept *person* is to specify the persistence conditions without using the very concept *person*⁵⁰. To find out the extension of the concept *person*, what one has to do is to apply the concept of person in various epistemically possible thought experiments. An example would be the Gettier case. The concept *knowledge* was traditionally analysed as justified true beliefs. However, in the Gettier case, a subject has a justified true belief P but we have the intuition that she doesn't know that P. Such intuition is from the application of the concept of knowledge. "Mutatis mutandis, we can ask whether, in circumstances described in a certain way, without appeal to that concept, the concept of personal identity would have application" (1989/2003, p. 196). The concept of personal identity, that is, the concept of person when applied to cases concerning persistence is a substance sortal: if an entity x fails to be a member of a substance sortal, x ceases to exist. Contrasting phrase sortal: if an entity x fails to be a member of a phrase sortal, it would continue to exist as a member of another sortal concept⁵¹.

⁵⁰ "The hope is that such conditions can be informatively specified, i.e. can be specified without the use of the very concept K which is being analysed" 1989/2003 p.88

⁵¹ Being a kitten will be an example of a phrase sortal

The Transplant Intuition: The Transplant Intuition refers to the intuition we have regarding a brain transplant case in which A's brain is transferred to B's body with consequent psychological traits. The intuition is that when the surgery is done and the B-body person who has A's brain and psychological traits wakes up, we identify that person *is* A. In other words, we have the intuition that *the person goes where psychological life goes*. Presumably, the Transplant Intuition will be induced in any case that has the same structure, including the Brown/Brownson case by Shoemaker and the Prince/Cobbler case by Locke. Noonan suggests that the argument from the Transplant Intuition is the "foremost objection to the Biological Approach (1989/2003 p.192)" and "can perhaps be better appreciated if it is seen as the outcome of a typical philosophical thought experiment." (1989/2003, p.195)

The Indeterminacy Thesis: Noonan endorses that it is possible that one's personal identity over time is indeterminate, and calls this type of cases 'borderline case'. In particular, the borderline cases he identifies are the Brown/Brownson case and the Methuselah case. In a borderline case, "the statement of identity in question is determinately neither true nor false because of the vagueness inherent in our concept of personal identity, and consequently it is neither determinately true nor determinately false that the original person in the case still exists after the various events in it have occurred" (1989/2003 p. 105). Rejecting the metaphysical view on vagueness, he argues for the semantic view (p.110-5, 2013) on vagueness, and employs it in the course of explaining the indeterminacy in the borderline cases he identifies (p.115-9). In this view, indeterminacy in the truth-value of the identity statement arises from referential indeterminacy of one of the terms employed in the identity statement.

Noonan's main argument for his neo-Lockean account is thus largely based on the Transplant Intuition and that the concept of personal identity is structured around a set of necessary and sufficient conditions. He argues that the Transplant Intuition is something that we must accommodate in giving an adequate theory of personal identity over time. Since neither the attempt of explaining it nor that of explaining it away is successful⁵², the only reasonable thing to say is that the Transplant Intuition confirms that neo-Lockeanism is true, unless there is a special

⁵² Noonan 1989/2002 p.197-201

reason for casting doubts upon the method of cases or the structure of the concept of personal identity.

Reference

- Dummett, Michael A. E. (1981). *The Interpretation of Frege's Philosophy*. Harvard University Press.
- Frege, Gottlob (1956). The thought: A logical inquiry. *Mind* 65 (259):289-311.
- Lewis, David K. (1986). *On the Plurality of Worlds*. Blackwell.
- Nidditch, P. H. (ed.) (1978). *A Treatise of Human Nature*. Oxford University Press.
- Ninan, Dilip (2016). What is the Problem of De Se Attitudes? In Stephan Torre & Manuel Garcia-Carpintero (eds.), *About Oneself: De Se Thought and Communication*. Oxford University Press.
- Noonan, Harold W. (2013). In Defence of the Sensible Theory of Indeterminacy. *Metaphysica* 14 (2):239-252.
- Noonan, Harold W. (2010a). Personal identity and its perplexities. In Georg Grasser and Matthias Stefan (eds.), *Personal Identity: Complex or Simple?*. Cambridge University Press.
- Noonan, Harold W. (2010b). Persons, animals, and human beings. In Joseph Keim Campbell, Michael O'Rourke & Harry Silverstein (eds.), *Time and Identity*. MIT Press.
- Noonan, Harold W. (2010c). The thinking animal problem and personal pronoun revisionism. *Analysis* 70 (1):93-98.
- Noonan, Harold W. (2003). *Personal Identity (2nd edition)*. Routledge.
- Noonan, Harold W. (2001). The epistemological problem of relativism – reply to Olson. *Philosophical Studies* 104 (3):323 - 336.
- Noonan, Harold W. (1998). Animalism versus lockeanism: A current controversy. *Philosophical Quarterly* 48 (192):302-318.
- Olson, Eric T. (2007). What are we? *Journal of Consciousness Studies* 14 (s 5-6):37-55.
- Olson, Eric T. (2004). Animalism and the corpse problem. *Australasian Journal of Philosophy* 82 (2):265-74.
- Olson, Eric T. (2003). An argument for animalism. In John P. Lizza (ed.), *Defining the Beginning and End of Life: Readings on Personal Identity and Bioethics*. Johns Hopkins University Press.

- Olson, Eric T. (2002). Thinking Animals and the Reference of 'I'. *Philosophical Topics* 30 (1):189-207.
- Olson, Eric T. (1997). *The Human Animal: Personal Identity Without Psychology*. Oxford University Press.
- Perry, John (1979). The problem of the essential indexical. *Noûs* 13 (December):3-21.
- Quine, W. V. (1976). Worlds away. *Journal of Philosophy* 73 (22):859-863.
- Russell, Bertrand (1912). *The Problems of Philosophy*. Barnes & Noble.
- Russell, Bertrand (1910). Knowledge by acquaintance and knowledge by description. *Proceedings of the Aristotelian Society* 11 (5):108--28.
- Shoemaker, Sydney (1999). Self and body. *The Philosophers' Magazine* 73 (8):29-29.
- Shoemaker, Sydney (1999). Critical Notice. Eric Olson, the human animal (new York: Oxford university press, 1997). *Noûs* 33 (3):496–504.
- Shoemaker, Sydney (1998). Causal and metaphysical necessity. *Pacific Philosophical Quarterly* 79 (1):59–77.
- Shoemaker, Sydney (1984). *Personal Identity*. Blackwell.
- Shoemaker, Sydney (1963). *Self-Knowledge and Self-Identity*. Cornell University Press.
- Snowdon, Paul F. (1991). Personal identity and brain transplants. In David Cockburn (ed.), *Royal Institute of Philosophy Supplement*. New York: Cambridge University Press. pp. 109-126.
- Stalnaker, Robert C. (1999). *Context and Content: Essays on Intentionality in Speech and Thought*. Oxford University Press UK.

Non-exclusive licence to reproduce thesis and make thesis public

I, _____,
(author's name)

1. herewith grant the University of Tartu a free permit (non-exclusive licence) to:
 - 1.1. reproduce, for the purpose of preservation and making available to the public, including for addition to the DSpace digital archives until expiry of the term of validity of the copyright, and
 - 1.2. make available to the public via the web environment of the University of Tartu, including via the DSpace digital archives until expiry of the term of validity of the copyright,

_____,
(title of thesis)

supervised by _____,
(supervisor's name)

2. I am aware of the fact that the author retains these rights.
3. I certify that granting the non-exclusive licence does not infringe the intellectual property rights or rights arising from the Personal Data Protection Act.

Tartu/Tallinn/Narva/Pärnu/Viljandi, **dd.mm.yyyy**