National Research University HIGHER SCHOOL OF ECONOMICS

As a manuscript

Tourko Dmitry

### THE REALITY OF THE SELF IN CONTEMPORARY ANALYTIC PHILOSOPHY

**Dissertation summary** 

for the purpose of obtaining academic degree Doctor of Philosophy in Philosophy

Academic supervisor: Diana Gasparyan,

Professor, Chief Research Fellow,

PhD in Philosophy,

Candidate of Sciences in Philosophy

## DISSERTATION DESCRIPTION (ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ)

The thesis straddles general ontology — ontology proper – and its subfield, personal ontology. Personal ontology, as Eric Olson states in his seminal exploration of the topic, addresses the questions: "What are we? That is, what are we, metaphysically speaking? What are our most general and fundamental features? What is our most basic metaphysical nature?" (Olson, 2007, p. 3). Apart from that, Olson says, personal ontology deals with the questions like (1) What are we made of? Are we made of matter at all? (2) If so, what kinds of matter are we made of (are we brains, or whole bodies, or something else?) (3) What parts do we have (one, finite number, indeterminate number of part combinations?)<sup>1</sup>. Ontology proper is "the general study of how things are or can be or must be" (Strawson, 2009, p. 1). It is also interested in the question "What general kinds of entities exist in the world?".

Accordingly, the personal-ontological part of this thesis addresses the question "What are we, metaphysically speaking?". The general-ontological part thereof seeks to answer the question "Do we exist?" or "Are we real?". It is obvious that the two are connected: your answer to the first question is probably going to affect your answer to the second one. Therefore, the answer to the former question is expected to precede that to the latter. First we figure out our basic metaphysical nature – or at least what it seems to us. Then we try to answer whether what this idea of ourselves signifies has any place in nature. Do we exist on par with rocks, protons and galaxies? Or do we exist as something parasitic on some other basic things? Or rather something less (something more?) than those basic things? Or is our common idea of what we are, in the way we commonly

<sup>1</sup> The principal question of personal ontology (*What are we?*) is connected to, but distinct from the question *How do we persist over time?* — the question of diachronic personal identity. It does not matter much if we consider personal ontology part of a wider personal identity philosophy or a distinct topic in its own right. As such, both personal ontology and personal identity straddle metaphysics and philosophy of mind. Although this thesis is concerned with personal ontology, I do not entirely ignore diachronic personal identity either.

conceive this idea, something entirely fictitious and illusory? Someone who answers the latter question affirmatively – an anti-realist about our own existence – might instead reformulate the first question in the following way: "What is it that we *seem* to be?". Their answer should strive to elucidate why this idea is fictional and has no merit in reality: why we are not real at least as the things we suppose ourselves to be. So for both realists and anti-realists it seems the personal-ontological problem takes precedence. First, say what you think we are; then try to figure out if we are real.

The term *we* in "What are we?" is ambiguous and equivocal. For now, I will take *we* to be synonymous with persons or selves<sup>2</sup>. So, the two questions asked about *me* (as part of *us*) are identical to those asked about the self which is I. Then the two principal questions of this thesis can be expressed as "What are persons, selves, or subjects?", and "Are those real?". We can call these two *the nature question* and *the reality question* of personal ontology, respectively. The nature question is related to but distinct from the question of the conditions for selfhood: What are the necessary and sufficient conditions that *x* must satisfy in order for *x* to be a self or to possess selfhood?

The nature question of personal ontology can be posed both in the plural and in the singular. So far, following Olson, I have been asking it in the plural: "What are we?", i.e. "Which kind of things do I belong to along with beings similar to me?". The nature question can also be formulated in the singular: "What am I *specifically*?", or "What differentiates me from beings who I share a kind with?". Put this way, it is clearly a different question. Even if I know the answer to the plural question, and thus know a great deal about my metaphysical nature as a member of a certain kind, I can still be ignorant regarding the answer to the singular question – regarding my nature as a particular as opposed to the nature of

<sup>2</sup> I treat the terms *person* and *self* (but not *subject*) interchangeably. In this synonymy, I follow early modern classics: both Locke (Locke, 1975, p. 346) [*Essay* II.26.24] and Hume (Hume, 1978, p. 173)) [*Treatise* 1.4.6.1] apparently used self and person as equivalent terms.

my kind. The most of this thesis is meant to be a search for the answer to the plural question. However, ch. 7 also approaches the singular question.

As a matter of fact, the lack of a common conceptual language and methodology dominates personal ontology and philosophy dealing with selfhood and subjectivity in general. The reason for this is not the field's inchoate state, but something else. The relevance and novelty of this work comes from its eagerness to recognize there are at least three distinct perspectives, levels of analysis or modes of description associated with the philosophical study of the self. These are the approaches, or groups of approaches, I prefer to call science-first (which, when making claims about the self, prioritize data from neuroscience, cognitive science and, to a lesser extent, evolutionary psychology), *metaphysics-first* (which put an emphasis on traditional arguments and methods of speculative metaphysics) and experience-first (which favour phenomenological<sup>3</sup> analysis of our experience and sense of self) modes of description of the self. I do not want to say that these three perspectives are clear-cut, always distinct and always incompatible. In fact, I make the opposite point. Obviously, the results of someone's phenomenological investigations of experience can have huge significance for their metaphysical views (as is the case with Galen Strawson). Any science-oriented philosophical study of the mind can make perfect use of someone's phenomenological reports ("third-person phenomenology", as Daniel Dennett calls it, is of course admissible in data-oriented research of mind and consciousness). Neurophilosophy, a kind of science-first approaches, is also implicitly metaphysical because it seeks to revise folk-beliefs about personal mereology (as is evident in Patricia Churchland's insistence that we should resist "the temptation to think of the self as a singular entity" (P. S. Churchland, 2002b, pp. 308-309)).

<sup>3</sup> By phenomenology, I do not strictly mean the historical tradition in mostly European philosophy closely associated with Edmund Husserl, his students and his legacy. Rather, I mean the broad method of philosophical investigation of consciousness which always has to start from investigating conscious experience. As such, it may be rooted in and closely connected to Husserlian and post-Husserlian phenomenology, but also can be independent of it. By metaphysics about persons, I mean speculative reasoning about their fundamental nature, reality, mereology, supervenience, ontological category, persistence and relations.

But the obvious difficulty stemming from this plurality of descriptions is explanatory incongruity that inevitably arises when we juxtapose the results of science-first, metaphysics-first and experience-first descriptions of the self. As an example, consider the unity of the self, a key element in Strawson's characterization of our own day-to-day experience (Strawson, 2009, p. 3). It seems true that, as he claims, we naturally have a strong sense of ourselves as persistent and unified mental entities or subjects. If we accept the basic presupposition of neuroscience that everything in the mind must first be in the brain, there must be a neural correlate for every mental capacity. So there must be a brain region or a network of neuronal connections that is responsible for cognitive activity associated with the unified self. A review by Gillihan and Farah (Gillihan & Farah, 2005) indicates that such a region or network does not exist; there is no unified self in the brain. The self, in the words of Vogeley and Gallagher (Vogeley & Gallagher, 2011), is "everywhere and nowhere" in the brain. So our experience tells us we are whole unities, while science tells us we are nothing of the sort. Soul-theory metaphysics can endorse the former view, while the metaphysics of mereological nihilism can rally for the latter one. No matter which you prefer, it will instantly put you at odds with either experience-first ("phenomenological") or science-first description of the self. Similarly, our sense of self has a strong embedded feeling of presence from which we naturally deduce the "fact" that we exist. Both Metzinger's neurophilosophical eliminativism (his phenomenal selfmodel theory, SMT) and Buddhist Reductionist could not disagree more. I return to these and similar conflicts of descriptions throughout the thesis, especially in ch. 3.

The question whether such warring intuitions, arguments, moves and descriptions of the self stemming from the three modes of description can be aligned or reconciled, or whether they are fundamentally untranslatable and self-contained, remains unresolved. However, throughout this thesis I will try to correlate descriptions of all the three levels, noting the difficulties that come up along the way. I will show that some problems imminently arise when we try to

reduce one mode of description to another. My general thesis here is that while multiple descriptions are not mutually reducible and not always compatible, they can have points of convergence — or a shared root from which they grow. Recognizing such points has the benefit of facilitating the dialogue between various research programmes and their practitioners. I argue that in the case of science-first, experience-first and metaphysics-first descriptions we can start looking for points of convergence in the idea of *reflexivity*, by which I mean the ability of some things - some systems, perhaps exclusively living systems - to represent their own states for themselves in a synthetic and epistemically closed manner. Reflexivity is the *for*-modality of some internal states, including mental states; and I believe we can make perfect sense of this modality from all of the three perspectives. In biological systems, reflexivity is the necessary condition for sentience: some living organisms which have developed reflexive systems we might as well call just sentient beings. Reflexivity helps us establish the first uncontroversial point from which we can further flesh out the concept of the self as an experiential entity.

Another reason for this thesis' relevance is the significance and urgency of various questions in metaphysics of persons for so many "real-life" practical domains. Personal ontology has a lot to offer to applied problems today, as I hope to show in the conclusion. Bioethics, animal ethics, environmental ethics, AI ethics, legal and moral personhood – a lot in these fields hinges on fundamental ideas and concepts supplied by the philosophical discourse on persons. Many of the listed fields are becoming increasingly prominent nowadays for a variety of reasons. This is why we need more philosophical work on foundational concepts and problems, specifically on the key concept of a person. Only when we figure out the fundamentals can we hope to achieve the conceptual clarity required to solve problems in applied fields. So it is important to get clear on the theory first before we get clear on the issues.

The final reason for this thesis' relevance is the general opportuneness of personal ontology today, perhaps amplified by the aforementioned demand coming from applied fields. Unlike the problem of diachronic personal identity, the topic of personal ontology was for the most part neglected in Analytic philosophy, until recently. It is not anymore – perhaps because by the late 20<sup>th</sup> century, Analytic philosophy greatly diversified and expanded thematically (or as some say, died). So it would have been impossible not to re-discover the classical problem of what we are. The views that were non-existent or poorly articulated previously animalism, narrativism, phenomenalism, neurophilosophical anti-realism and Analytical Buddhism – have sprung up over the past several decades. Now the topic is well-published on (see Literature review for a more detailed discussion of these matters). Concerning the relation between the problems of diachronic personal identity and personal ontology, I share David Wiggins' view he calls sortalism: "the position which insists that, if the question is whether a and b are the same, it has to be asked what are they?" (Wiggins, 2012, p. 1), also endorsed by Igor Gasparov: "Without an adequate answer to the question 'What am I?' it is impossible to answer the question of what constitutes personal identity" (Гаспаров, 2021, p. 78). In order to investigate the persistence conditions of persons, we need first to get clear on what persons are. In a word, personal ontology must precede diachronic personal identity - this is why the current resurgence of interest in personal ontology is long overdue.

#### Literature review (Степень разработанности темы исследования)

The core questions of personal ontology – the nature and reality of the self, person and subject – are also among the core metaphysical questions; and as such they have been debated since the dawn of philosophy, albeit not necessarily in these particular terms. It would be impossible to cover here all relevant ideas about the self and person in philosophy, even in Western philosophy. Barresi and Martin's *The Rise and Fall of Soul and Self: An Intellectual History of Personal*  *Identity* (2006) (Martin & Barresi, 2006) offers an increasingly broad historical account of how our thinking about our metaphysical essence has developed over millennia.

*Early modern philosophy. Descartes, Locke, Hume.* Among pre-20<sup>th</sup> century sources, this thesis only discusses the milestones of early modern philosophy relevant to the topic. The most important of those are Descartes' Meditations on First Philosophy (1641), especially the second meditation; Locke's An Essay Concerning Human Understanding (1689) (specifically, Book II, Ch. XXVII Of Identity and Diversity) and Hume's A Treatise on Human Nature (1739) (specifically, Book I Ch. IV, Section 6 Of Personal Identity). The account I am going to defend, ExPR, draws a huge deal from early modern philosophy. It is a sort of Neo-Cartesianism because it agrees with Descartes that we are fundamentally thinking things. It is in a sense Neo-Lockeanism because it shares Locke's idea that personal identity over time consists in the sameness of consciousness. It is a kind of Neo-Humeanism because it envisages persons as fundamentally nothing more than rapids of "impressions" – experiential processes. In an important sense – which I hope will become evident as I rally early modern material to my cause further down the road – Descartes, Locke and Hume were phenomenalists; specifically in the sense that they connected our existence and fundamental metaphysical nature with our thinking, consciousness and flux of impressions – or, if I am forgiven such an anachronistic generalization – with our having *conscious experience*.

*Contemporary theories of personal ontology*. Broadly, many if not most contemporary theories of both personal identity and ontology can be divided into "brute physical" and mentalistic accounts. The representatives of brute physical accounts are: the view are that we are living bodies (Thomson, 1997), the view that we are "highly complex, composite material objects" (Hudson, 2001, p. 1), animalism – the view that each of us is identical to a human animal, the view that

we are brains or brain parts (more on these two below) as well as all naturalistic and neurophilosophical – roughly, science-first – conceptions. All these accounts take both our diachronic continuity and basic metaphysical nature to inhere in some material, non-mental properties and facts. Mentalistic accounts on the other hand (most importantly, Neo-Lockean psychological theories and Neo-Lockean consciousness-first theories or phenomenalism) claim our continuity and nature have to do with – or primarily consist in – our mental properties and contents. Note that this distinction does not correlate with the physicalism-antiphysicalism divide in the mind-body problem: many proponents of mentalistic accounts are also physicalists in their general metaphysics. That said, many contemporary theories transcend the brute physical-mentalistic divide because the elements that the nature of persons consists in are, according to these theories, can be both mental and physical (or neither). As opposed to these "neither-nor" theories, this thesis defends a species of the purely mentalistic class of theories – at the same time operating on the underlying physicalist assumption that mental entities, events and states are also *broadly* physical.

Animalism. The one account of personal ontology that has been on the rise in the two recent decades is animalism. In current debates, animalism seems to be at the intellectual forefront of physicalistic accounts as it has mostly superseded other "brute physical" views in personal ontology and identity, such as the old view that each of us is identical to a physical body. Animalism is the idea that each of us is numerically identical to an animal of the species *H.s.* There is a certain animal organism of a certain species, and you and that organism are the same. Animalist philosophers<sup>4</sup> are generally uninterested (Olson, 2007, p. 44) in the nature question of selves and persons. At the same time, they insist only that "we are animals, not

<sup>4</sup> Bailey (Bailey, 2015, 2016, 2017; Bailey et al., 2021; Bailey & Elswyk, 2021); Blatti (Blatti, 2012, 2014); Geddes (Geddes, 2013); Olson (Olson, 1999, 2003, 2015); Snowdon (P. Snowdon, 2009; P. F. Snowdon, 1995, 2004, 2014) and Toner (Toner, 2011). Direct critics include Duncan (Duncan, 2021), Hudson (Hudson, 2007), Johnston (Johnston, 2007, 2016), Shoemaker (D. Shoemaker, 2016) and Unger (Unger, 2000) among many others. For a collection of papers, see (Blatti & Snowdon, 2016). For an excellent overview of animalist arguments and positions, see (HexaeB, 2021).

that people in general are; so it is compatible with the existence of people who are not animals (gods or angels, say), and of animals – even human animals – that are not people" (Olson, 2007, p. 24). So the extent to which *personal ontology* is associated with *human animal ontology*, is unclear, and I will devote some time to discuss animalist-personalist conceptual contradictions in ch. 6.

Brain views. As opposed to whole animals, some "brute physical" and science-first theorists identify the self or simply us with the conscious and thinking brain or its parts. The intuition that we are brains or brain parts is heavily supported by the effects of corpus callosotomy on personal identity. These effects are interpreted by some as the single brain supporting two or more numerically distinct persons (see (Puccetti, 1973) for one of the earliest accounts; see (Rigterink, 1980) for a critique); or they are interpreted in the way that suggests "there is no whole number of individual minds that [split brain] patients can be said to have" (Nagel, 1971, p. 409), and the "simple idea of a single person will come to seem quaint some day" (Nagel, 1971, p. 411). The brain view is supported by a family of scenarios: Shoemaker's brain transplant (S. Shoemaker, 1963, pp. 23–24), Parfit's surviving head and surviving cerebrum (Parfit, 2012, p. 11), and Johnston's gruesome guillotining (Johnston, 2016, p. 113). All these scenarios are directed against animalism or the bodily view (see (Olson, 1999, pp. 114-119) for an animalist rejoinder) as they are meant to summon the intuition that personal identity is preserved by preserving (living) heads and brains and *not* the rest of the body or animal which is not essential for the purposes of survival and identity.

*Realist theories. Realists* about the self<sup>5</sup> are theorists who explicitly affirm the existence of the self as some real, concrete, non-conventional and non-constructed entity, albeit their understanding of that entity is extremely varied. Unlike animalists and brain theorists, philosophers with an explicitly realist

<sup>5</sup> The specific terms *anti-realism* and *realism about the self (self-realism)* are used by Krueger (Krueger, 2010), Albahari (Albahari, 2010), Metzinger (Metzinger, 2011), Tekin (Tekin, 2015) and Jennings (Jennings, 2020).

position usually conceptualize the self as an entity distinct from both animals and brains; as one would expect, their counterparts anti-realists deny the existence of such an entity. Let us now consider some realist and then anti-realist schools.

Soul theories, idealism and Neo-Cartesianism. There are a few dualist and idealist contemporary philosophers who defend the view that the self, taken to be our essence, is in fact our soul. Others defend Cartesian substance dualism with its thesis that we are minds or mental substances really distinct from our bodies or anything physical whatsoever. We can classify John Foster (Foster, 1991), David Lund (Lund, 1994, 2005, 2009), Colin McGinn (McGinn, 1996, p. 161), J.P. Moreland (Moreland & Rae, 2000; Moreland, 2014) and Richard Swinburne (Swinburne, 1986, 2013, 2019) as belonging to this category. Weir (Weir, 2023), also a soul-theorist, defends the view that property dualism directly implies substance dualism. (S. Shoemaker & Swinburne, 1984) is a classical debate between a dualist (Swinburne) and a physicalist who follows the psychological criterion of personal identity (Shoemaker). See (Corcoran, 2001) for a collection of papers by proponents and detractors of idealism and Neo-Cartesianism; see (Inwagen & Zimmerman, 2007) for a collection of contributions by selfproclaimed idealists, dualists and materialists with a section of personhood in Christianity.

*Phenomenalist or experience-first theories.* The most recognizable realist camp nowadays are the theorists who can be collectively dubbed *experientialists* or *phenomenalists* about the self — although that is almost never a self-designation. Phenomenalists are philosophers who connect the reality of the self with the reality of experience (Dainton & Bayne (Dainton & Bayne, 2005); Dainton (Dainton, 2008, 2012); Gallagher & Zahavi (Gallagher & Zahavi, 2020); Strawson (Strawson, 1997, 2009, 2011, 2017)), or with the *subjective character (mineness)* of experience (Zahavi, 2005).

Narrative theories. Another active and distinct group of accounts in personal ontology and identity is narrativism. Narrative theorists characterize selves and persons in terms of their capacity to create and interpret autobiographical narratives ((MacIntyre, 2007, Chapter 15), (Taylor, 1989, Chapter 2), (Carr, 1991), (Ricoeur, 1994, Chapters 5-6), (Schechtman, 1996, 2007, 2014); (Christman, 2004); (Davenport, 2012), (Rudd, 2009, 2012)). It is unclear where narrative theories are situated in the realist-antirealist divide. Narrativists typically see the self as an acting, speaking, responsible and story-telling entity (Ricoeur, 1994, p. 297) "whose unity resides in the unity of a narrative which links birth to life to death as narrative beginning to middle to end" (MacIntyre, 2007, p. 205). This characterization implies at least some ontological commitment to the existence of the self. Perhaps the general narrativist stance on the reality question is deftly expressed by Rudd: "The self is not something that just exists, and is then narrated (by itself or by others); it only comes to exist through its being narrated" (Rudd, 2012, p. 1). It follows that while narrativists never attempt to reify the self and think that its existence is predicated on the existence of a narrative, it is rare for them to embrace explicit anti-realism, i.e. claim that the self does not exist, is an illusion, a philosopher's myth, or is fully reducible to some ontologically primal elements. However, some narrativists espouse views quite in line with anti-realism.

Anti-realist theories. As opposed to realists, let us now consider "negativists" about persons. We can call the idea that the self does not exist, is an illusion, a residue of Cartesian dualism, a conceptual muddle, a piece of "philosopher's nonsense" (Anthony Kenny (Kenny, 1988)), or really is some other thing or bunch or things, *anti-realism about the self*. Anti-realism is also called the *no-self view* or *no-self theory* ((Siderits et al., 2010, pp. 4–5); (McClelland, 2019, p. 22)) or *non-entity theory* (Lowe, 1991, p. 84). In today's literature, anti-realists come in many shapes and colours and are motivated by a variety of reasons in their hostility towards the concept. A distinctive mark of contemporary anti-realism, however, is that it almost always denies the self as it is traditionally conceived – as

a unified, simple, single and persistent entity, i.e. as a substantial entity; but at the same time anti-realist theorists rarely consider other conceptions of the self, such as proposed by the phenomenalists just discussed (more on that in ch. 3 and 6). Let us now consider some varieties of antirealism.

Neurophilosophical and cognitivist anti-realism. The first camp of antirealists draw their arguments and inspiration from cognitive science and neuroscience. It is represented by Patricia Churchland (P. Churchland, 2013; P. S. Churchland, 2002b, 2002a, 2011); to a lesser extent Daniel Dennett (D. Dennett, 1992; D. C. Dennett, 1989, 1991); Bruce Hood (Hood, 2012); Tom McClelland (McClelland, 2019); and especially Thomas Metzinger (Metzinger, 2003, 2007, 2010b, 2010a, 2011). The general stance on the self here is reductionist or eliminativist. It is argued the self is nothing like what we suppose it to be, and our natural intuitions as well as any insights from the so-called "self-experience" should be abandoned. Thus, Patricia Churchland interprets the self as a bunch of the brain's representational capacities. She suggests recasting multiple and diverse meanings of the term *self* — multiple selves — as multiple brain functions. (P. S. Churchland, 2002b, pp. 308–309). Accordingly, ch. 3 of this thesis is an extensive argument against neurophilosophical anti-realism, especially in its Metzingerian iteration.

*Buddhist anti-realism*. The second large anti-realist camp actualizes arguments from Buddhism and other Indian philosophies ((Albahari, 2006, 2010, 2014); (Chadha, 2017, 2017, 2018, 2019, 2021; Chadha & Nichols, 2022); (Dreyfus, 2010); (Deikman, 1996); (Krueger, 2010); (Bihan, 2019); (M. MacKenzie, 2010, 2022; M. D. MacKenzie, 2007); (Siderits, 2010, 2011, 2017, 2019)).

The book I will focus on in my critique of Buddhist anti-realism is Mark Siderits' *Personal Identity and Buddhist Philosophy: Empty Persons* (Siderits, 2017). which reconstructs and defends the view Siderits calls Buddhist Reductionism. According to Buddhist Reductionists, persons are conventional, transient entities which do not exist as far as the ultimate reality is concerned. I engage with Buddhist Reductionism in ch. 5 of this thesis. As I demonstrate, core Buddhist Reductionist arguments hang on mereological nihilism, a view that composite things do not exist. So ch. 5 also formulates, analyzes and attacks mereological nihilism about persons.

#### Subject of inquiry (Объект и предмет исследования)

The subject of this thesis' inquiry is the referent or referents of philosophical terms *self, person* and *subject*. The two primary investigated subject matters are, first, the reality of their referent or referents (or: whether the terms are referring and not empty), and second, the nature of those. Correspondingly, the two principal research questions of the thesis are "What are persons, selves, and subjects?", and "Are those real?". It is opportune to call these two *the nature question* and *the reality question* of personal ontology, respectively.

#### Objectives (Цель и задачи исследования)

My principal objective in this thesis is to defend *experiential process-realism* (ExPR) as a theory of personal ontology – a theory about selves, persons and subjects. To achieve it, I aim for the following specific goals:

- draw all the conceptual distinctions necessary for a clear investigation of the concept *self* (chapter 1);
- provide a succinct naturalistic account of sentience and experiential selfhood (chapter 2);
- offer a detailed refutation of neurophilosophical anti-realism about the self (chapter 3);
- analyze self-experience and the problem of the subject of experience; defend a biconditional existential dependence view about the subject and its experience (chapter 4);

- argue against Buddhist anti-realism (Buddhist Reductionism) about selves and persons (chapter 5);
- supply direct arguments for experiential process-realism about the self as well as objections against rival accounts (chapter 6);
- 7) defend the view that persons or selves are irreducible entities (ch. 7).

In the Conclusion, I hope to show the relevance of the defended framework for some applied fields. I believe it is of paramount importance to continue our investigation of the self under various modes of description – science-first, experience-first (phenomenological) and metaphysics-first descriptions. This is why my underlying conciliatory purpose in this thesis is to show that different levels of description of the self are at least partially compatible, even if often mutually contradictory. I intend to offer the concept of reflexivity as the basic conception which can serve as the initial starting point for these three principal descriptions.

#### Methodology (Методологическая основа исследования)

1. The Method? Personal identity is a somewhat infamous topic in metaphysics because of its theorists' commitment to "fanciful", unbelievable and far-fetched thought experiments, often as if lifted from works of science-fiction. (This way of philosophizing about the self is simply called "the Method" by Dainton and Bayne (Dainton & Bayne, 2005)). Apart from the already mentioned Shoemaker's brain transplant and its variations, some of the most widely known experiments were developed by Derek Parfit. These are fission (person-branching by implanting identical hemispheres of the same brain into two different brainless bodies) (Parfit, 1971) and simple teleportation (recording all data about a person's body on Earth, destruction of the original body and then reassembly of a qualitatively identical body on Mars) (Parfit, 1984, p. 5); (Parfit, 2012, p. 199). Such over-reliance on imaginary scenarios has been roundly criticized. In this thesis, thought experiments will be used very sparingly and mostly to illustrate and

support arguments. In ch. 6 (*The survivability test*), I offer three scenarios in support of the essentiality of phenomenal property for persons.

2. The three descriptions. As already been noted several times, this thesis assumes there are three principal modes of description, approaches or viewpoints from which current theorists typically investigate the self: *science-first*, *experience-first* (phenomenological) and *metaphysics-first* descriptions. These descriptions differ in which epistemic sources they prioritize when they endeavour to philosophize about the self. The science-first description favours the data of brain and cognitive sciences and, to a lesser extent, evolutionary psychology and biology; the experience-first method seeks to understand the self by examining our phenomenal experience; while the metaphysics-first approach utilizes speculative philosophical arguments without foregrounding either scientific data or phenomenological insights. Although they are not entirely isolated from each other, they are not perfectly compatible either. Methodologically, it is helpful to keep this multiplicity of descriptions in mind when philosophizing about the self; because sometimes our commitment to just one of them prevents us from achieving the best results, as I hope to show in the thesis.

3. Modesty. The next methodological principle is inferring from the least premises and metaphysical parsimony – or perhaps modesty. The position I am going to defend (ExPR) will be derived from a minimum of premises in order to achieve the results that would be acceptable or at least understandable to representatives of all considered modes of description. Thus, I do not postulate any characteristics of the self at the beginning, but instead proceed from an understanding that I hope would cause little controversy. Such an approach is adopted in hopes of reconciling the (proponents) of the three levels of description mentioned above.

4. On guard against polysemy and ambivalence. It will also be important to mention several methodological dangers connected to our topic. The first of these

is conceptual ambivalence (e.g. see the section *Person, ego, subject, I and self* in ch. 1) to avoid which I strive to employ clear definitions and conceptual analysis. Different approaches to the self, its different aspects, definitions and understandings are not necessarily mutually exclusive, especially in an interdisciplinary context. However, semantic vagueness and overladenness of the concept prevent clarity, so in this text I try to reduce ambivalence through analytical work.

5. Strong pluralism and the challenge of conventionalism. The next conceptual and methodological pitfall after polysemy and ambivalence is pluralism, that is, the reification of the multidimensionality described above. It is difficult to deny that a person — at least a human one — includes several layers, levels and aspects. Each of us is a multitude of things, properties and dispositions at the same time. Such a *moderate version of pluralism* is unlikely to raise many objections. But one can also imagine a strong pluralism that asserts that all aspects of the self are equal in importance, and that there is no one aspect that would have a decisive metaphysical significance - say, would be essential for personal survival. As Braddon-Mitchell and Miller's work indicates (Braddon-Mitchell & Miller, 2004, 2020), such pluralism is also connected to *conventionalism*, the idea that the answer to the question What is the self? is not a true proposition, but only a judgment reflecting consensus, value-preferences or practical concerns. For an argument against conventionalism (and in favor of what he calls "realism"), see (Merricks, 2001). This thesis adopts Merricks' realism as I argue against conventionalism in ch. 2.

6. Essentialism and qualified pluralism. We are many things – both literally (we are probably not simple indivisible substances) and conceptually (there are many layers to us). But I am not really interested in positing plurality; such a thesis would be rather trivial. Instead, *qualified pluralism* is going to be my methodological principle. I want to look for the fundamental layer – the minimal or

essential self; the foundation of the complex person-structure upon which any higher-order stories have to rest on. Qualified pluralism admits we are many things, but among those things there is one thing (or more) which is sine qua non. Thus, I reject strong pluralism, accept qualified pluralism and offer arguments in defence of essentialism: the idea that there is a fundamental and necessary level of self on which all others are built.

#### Novelty (Научная новизна исследования)

The thesis's novelty consists in the following theoretical achievements:

—a new framework (ExPR) combining realism, phenomenalism and process ontology of selves is developed, described and defended;

—the self is uniquely considered from the three seemingly contradictory perspectives: science-first, metaphysics-first and experience-first approaches;

—an attempt to reconcile the three perspectives is made via the concept of reflexivity;

—the self and the subject of experience are considered as distinct entities; so separate accounts of their reality and nature are offered;

—a novel approach to the basic ontological category of selves is offered: the self is considered consecutively as an entity, property and process;

—a new light is cast on the established distinctions, concepts and arguments in philosophy of mind, personal identity and ontology: the primitive argument (ch. 1), the processual cogito (ch. 6), the survivability test / divergence scenarios (ch. 6);

—numerous new distinctions, concepts and arguments are made: the intuition test for a conception of selfhood (ch. 1), the naturalistic argument (ch. 2), the experience-properties as self-properties test (ch. 4), the phenomenal dependence argument (ch. 2), the presentist argument for transientism (ch. 4), the

18

no ontologically independent things but atoms argument (ch. 6), the value-only argument for person-priority (ch. 6), the knowledge argument for irreducible persons (ch. 7) and others.

#### Positions (Основные положения, выносимые на защиту)

I call the view I am going to argue for in this thesis *experiential process-realism* (ExPR). ExPR is a *moderate* (non-substantivist), *naturalistic*, *egological intraneous*, *non-reductive*, *further-fact*, *emergentist*, *processual* realism about selves. The seven chapters to follow are a thorough defence of ExPR and associated positions, and a critique of anti-realist views, specifically those of contemporary *neurophilosophical anti-realists*, *Buddhist Reductionists* and *mereological nihilists*. As such, ExPR provides answers to most of the core questions in personal ontology, including the nature question (the theses 1-2 below) and the reality question (3-4) — but also many others like the nature of subjects of experience (5), the mereological nature of persons (6), the person's basic ontological category (7), the person's relation to its ontological base (8, 9) and its reducibility (12). So let me list all these answers as positions I am going to defend throughout this work.

#### 1. Anti-realism about selves is false.

If anti-realism about selves means propositions [I am or have a self] or [Selves exist] or [There are actually existing persons] are meaningless or false, anti-realism is false. You and I are real – provided you and I are persons. If the reader finds this thesis trivial, the recent philosophical (especially neurophilosophical and neo-Buddhist) pushback against the alleged reality of selves begs to differ.

Refutations of anti-realism: ch. 3 (neurophilosophical anti-realism) and ch. 5 (Buddhist Reductionism and mereological nihilism).

2. Selves are real.

Consequently, in regard to the reality question, my realism opposes all kinds of anti-realism about the self: any views that argue selves or persons are illusory or do not exist. They are real as much as other mundane and natural physical objects are real. Sentience, experientiality or selfhood is concretely realized property of, or a concrete process in some naturally evolved systems. There are sentient things in the world. This fact is part and parcel of a full physicalist description of the world. So those who believe there are no selves are mistaken: at least on such a naturalist construal, selves exist. In a word, I agree with Dainton who takes "conscious states to be just as real, just as much parts of concrete reality, as protons and electrons, or stars and planets" (Dainton, 2008, p. xiv).

Arguments: the primitive argument (ch. 1), from history to reality principle, the naturalistic argument (ch. 2), the processual cogito (ch. 6).

#### 3. Selves are essentially experiencers.

On the nature question of personal ontology (*What are we as persons* essentially?), I side with consciousness-first or experience-first theorists of recent years: Galen Strawson, Dan Zahavi and especially Barry Dainton. In personal ontology, I argue for an account of selfhood similar to Dainton's *C*-theory. According to Dainton, we are *C*-systems: bunches of experiential capacities. I argue for a closely related, but a less ambitious claim: every person is identical to an experiential entity, or experiencer. Thus, the second central thesis of this work (after realism) is that selves, or persons, are experiencers — experiential entities. You, I, or any other actual or possible person, be it human, non-human animal, alien, angel, conscious machine or other, are identical to a numerically distinct experiencer or thinking thing. That is, the best way to handle selves or persons is to conceive them as experiential entities. An experiential entity is any thing which in virtue of its neuronal or other organization is able to support phenomenal experience; its

experientiality being dependent on supporting substratum (base), whatever its nature. Speaking in terms of properties, if my selfhood is my essential property and selfhood is best understood as experientiality, then my essential property is experientiality.

Combined, theses 2 and 3 form what I prefer to call *experiential realism* about selves. I argue that sentience and experientiality are identical properties. Therefore, experiential selves are identical to sentient entities.

Arguments: the annihilation argument, the intuition test for a conception of selfhood (ch. 1), experience-properties as self-properties (ch. 4), the phenomenal dependence argument (ch. 2), the value argument for experientialism (ch. 6), the survival test / divergence scenarios (ch. 6).

4. Selves are units of nature.

My experientialist realist account is reinforced by a principal thesis: experience is a proper part of the natural world. It is a property of things that came into being gradually, spontaneously and as a useful adaptation through evolution and natural selection. It is neither part of the supernatural domain (as a strong realist would often claim), nor an abstract and formal entity (as a narrativist would claim), nor a conventional entity.

Argument: the primitive argument (ch. 1); the naturalistic argument, selves as natural kinds (ch. 2).

#### 5. The egological intraneous view about the subject of experience is true.

Apart from offering answers to the nature and reality questions, my additional task is to clarify the nature of subjects of experience (which I take to be distinct from persons and selves). I argue that experientiality comes with a built-in subjectivity: every experience is intrinsically subjectival. Minimal subjectivity is embedded in every experience, for any kind of person. However, I claim that the subject is necessarily an experiential entity itself, meaning it has no existence outside of its experience. Thus, subjects of experience are not "independent" objects or entities either. In other words, strong egological accounts of the subject which conceive it as standing above and beyond its experience, are implausible. Instead, subjecthood is a built-in, necessary, inherent property of any experience. I explore several problems connected to the subject of experience and its several conceptions in ch. 4.

Argument: the presentist argument for transientism (ch. 4).

6. Reductionism and mereological nihilism about persons are false.

Reductionism about x is a view that x is *really just* y. Our taking x to be anything but y is a result of our epistemic limitations. While reductionism is arguably compatible with realism, in the realist-anti-realist debate it explicitly gravitates towards the latter. Mereological nihilism is a view that composite things do not exist. Mereological nihilism about any composite x is a form of anti-realism about x. I discuss and argue against both reductionism and mereological nihilism in ch. 5, in connection with their Buddhist iterations as expounded by Mark Siderits. My argument is made mostly on personalist, functionalist, and emergentist grounds.

Arguments: see ch. 5.

7. Selves are best understood as experiential processes.

Next, I focus on the question of relation between experientiality and its ontological base, the elements of which it existentially depends on. In a very broad sense, a base of x is an entity, substance or property whereon x ontologically depends, supervenes or emerges<sup>6</sup>. I claim selfhood or

<sup>6</sup> I prefer to explain the self-base ontological dependence relation in terms of generic existential necessary dependence: "x is generically existentially necessarily dependent on  $Fs \ll df$ . Necessarily, x exists only if some Fs exist" (Hoeltje et al., 2013, p. 40), Fs in our case being constitutive elements of the base (base parts). Contrast it with the rigid existential necessary dependence: "x is rigidly existentially necessarily dependent on  $y \ll df$ . Necessarily, x exists only if y exists" (Hoeltje et al., 2013, p. 40), Fs in our case being constitutive elements of the base (base parts). Contrast it with the rigid existential necessary dependence: "x is rigidly existentially necessarily dependent on  $y \ll df$ . Necessarily, x exists only if y exists" (Hoeltje et al., 2013, p. 38). Supervenience is fundamentally understood thus: "To say that so-and-so supervenes on

personhood is an ontologically based or dependent process. Persons are ontologically dependent in the sense that their base (such as, in some biological person-kinds like humans, their neural substratum) serves as a substratum for experientiality which is non-identical to the substratum. Experientiality is a property E of an organized system s such that if E(s) is true, s can support experience. Several considerations discussed in the thesis indicate that each of us, persons, is best understood as an uninterrupted capacity for experience, or even better – as an uninterrupted experiential activity or process, rather than a distinct object or substance. The reasons to accept processualism come from *the composite view* about persons (the idea that persons are not mereologically simple things), *the unimportance of base identity for personhood* view (see thesis 8 in this list) and the fact that processualism helps us solve some important problems in personal identity and ontology (see ch. 6, esp. sections *Self-processes: third approach (fire and fuel)*, *The bridge problem, The processual cogito*).

Argument: the Humean zoom-out (ch. 2); pragmatic and metaphysical considerations in ch. 6.

8. Selves are not necessarily anchored in their ontological bases: the sameness of base is not necessary for the sameness of person.

As persons, we are necessarily experiential entities; and experientiality is a based property; but we are also based, occurrent *processes* of experientiality. However, as I hope to show using recognizable scenarios at the end of ch. 6, selves are not necessarily rooted or anchored in some rigidly fixed base whose identity must be preserved for personal persistence to obtain. In other words, the self and its base elements are in a *generic existential necessary dependence* relation (see

such-and-such is to say that there can be no difference in respect of so-and-so without difference in respect of such-and-such" (Lewis, 1983, p. 358). I use the term *base* coextensively *base elements*, meaning the parts or elements on which a higher-order entity depends.

the note on the previous page for definitions). It is conceivable that the self maintains its numerical identity over time even if its base identity is reset. So base identity is distinct from personal identity. Above, I called this development of an old Lockean idea *the unimportance of base identity for personhood* view. This unimportance is what bestows metaphysical freedom on persons, as argued at the end of ch. 6.

Argument: see the discussion and scenarios in ch. 6.

9. Strong realism is false: selves are not ontologically independent.

We are metaphysically free from the identity of our bases, but not from the existence of our bases. Selves are not ontologically independent<sup>7</sup> entities with respect to some base that existentially supports and possibly subvenes them (if entities independent of any base are possible, perhaps only philosophical atoms and God would qualify). I argue that selves emerge and non-rigidly existentially depend on their bases. The group of views which deny that selves are ontologically dependent – I call it *strong realism* – is implausible.

Argument: the *no ontologically independent things but atoms* argument (ch. 6); also see the relevant discussion in ch. 6.

10. Non-phenomenalist theories of personal ontology are only accidentally true.

Argumentatively, my experientialist account of personal ontology opposes all "brute physical" and non-experiential "mentalist" accounts thereof. It implies that the views in personal ontology incompatible with experientialism are either false — just in case they are taken to be accounts of what persons *essentially* are; or true if they describe only some *accidental* properties of persons. I.e. they are true only if they do not claim to set

<sup>7 &</sup>quot;*A* is *ontologically independent* from *B*'s just in case *A* can exist without the *B*'s" (Hoeltje et al., 2013, p. 75).

necessary or minimal (non-experiential) properties of selfhood; or only if they speak about properties of just some persons, or only at some point of some persons' existence. This is why it would not be disparaging to call them *accidentally true*.

Arguments: the phenomenal dependence argument (ch. 2); see also ch. 6.

#### 11. Human nature is metaphysically indeterminate.

In accordance with the ontological recasting of selves as processes (thesis 7), selves are composites. That said, not only selves, but also unified human beings are composites. In part because human beings are composites, they have an indeterminate basic nature. There are several non-contradictory, complementary correct answers to the question *What are we?* (animals, thinking things, narrative-spinners etc.) if *we* is not laden with presuppositions. There is no factually correct answer to *What are we essentially?*, i.e. there is no fact of the matter which would force us to accept a certain answer. But our values and preferences more often than not incline us to think we are persons first and anything else (e.g. animals or bodies) second. But again, as I argue (see theses 2 and 10) there is only one correct answer to the question "What are we as *persons*, essentially?".

Argument: the value argument for person-priority (ch. 6).

12. The further-fact view and strong emergentism about persons are true; persons are irreducible entities.

The strongest and probably the most controversial claim of this thesis is that some properties of persons are strong emergent properties, i.e. they arise on a base of lower-order properties or elements which collectively lack the higher-order property. In connection to that, I argue that some facts about persons are further facts: they are epistemically isolated and not accessible from any other fact. A quasi-omniscient observer who knows all deducible and learnable facts would not be able to know further facts. I provide an involved argument for these two theses in ch. 7, where I conceptualize  $I^*$  (the occurrently conscious self) as opposed to "mere" persons and consider the question whether the quasi-omniscient observer (the Laplacian demon) would be able to infer I\*'s position in space and time. Eventually, I offer an argument that persons are ontologically irreducible things.

Argument: the knowledge argument for irreducible persons (ch. 7).

# Theoretical significance and practical applications (Теоретическая и практическая значимость исследования)

Apart from its successful arguing for the positions it argues for (if it indeed argues successfully), the thesis' theoretical significance consists in its attempt to comprehensively develop an ontology of selves: defend their processual understanding, conceive selves as based properties, investigate the nature of the subject of experience, defend a thorough anti-reductionist, realist and experientialist account. Additionally, the thesis' theoretical significance comes from its assumption that the three broad descriptions or perspectives on the self (science-first, metaphysics-first and experience-first), though often at odds, are not mutually exclusive or entirely incompatible.

The thesis' practical significance comes from potential applications of some of its core ideas to certain problems in ethics and adjacent fields. Namely, the thesis' results can have significant implications for the moral status of persons, artificial personhood and especially non-human animal personhood (see Conclusion for a brief discussion how), as well as for the philosophical treatment of death, survival and immortality.

#### Thesis structure (Структура диссертации)

In the Introduction, I review relevant literature and discuss limitations, objectives and methodology of the thesis.

Chapter 1 prepares the ground for the further investigation by making necessary conceptual distinctions. It attempts to approach and define the ambiguous term *self* from different perspectives. At the end of the chapter, I provide just a brief outline of the defended account of selfhood, leaving most of the argument for later chapters. In line with the tripartite thematic division into the science-first, metaphysics-first and experience-first philosophical descriptions of the self, apart from the Introduction and Chapter 1 the thesis is broken down into three respective blocks.

Chapters 2 and 3 represent the science-first block.

Chapter 2 endeavours to offer a succinct naturalistic account of selfhood which in that chapter is identified with sentience. The key thesis argued for in ch. 2 may be dubbed naturalistic realism: selves are best understood as natural and evolved entities.

Chapter 3 argues at length against neurophilosophical anti-realism about the self represented by Patricia Churchland and especially Thomas Metzinger. The first part of the chapter discusses the general gist and problems of neurophilosophical anti-realism. The second part of the chapter thoroughly engages with Metzinger's important book unequivocally titled *Being No One* (2003).

A single chapter 4 makes up the experience-first or phenomenological block, although some of its themes are also continued in chapter 5.

Chapter 4 considers the self from the phenomenological point of view. It tries to ostensively approach conscious experience, looks at the problem of self-experience, lists the kinds of self-experience in order to make it clear how they relate to the self and subject. Finally, it discusses core problems of the subject of experience, its nature and reality. The second part of the chapter makes a key argument for biconditional co-dependent yet distinct existence of the self and its subject.

27

Finally, chapters 5, 6 and 7 are included into the metaphysics-first block.

Chapter 5 argues against contemporary Buddhist anti-realism (Buddhist Reductionism) about the self and person as it is reconstructed and represented by Mark Siderits. It discusses the Buddha's own views from the Pali canon and the arguments against the existence of the self and person later Buddhist philosophers made. Finally, it argues against mereological nihilism, the view that composite things do not exist – the key premise of Buddhist Reductionism.

Chapter 6 directly tackles the self as a metaphysical entity. This is where the reader can find principal arguments for experiential process-realism about the self as well as objections against rival accounts, such as strong realism, the brain theory and animalism.

Finally, chapter 7 goes further in the purely metaphysical direction and asks whether persons are reducible to physical facts — and indeed to any learnable facts. It answers the question in the negative.

The conclusion discusses the results and briefly considers their applications for several applied fields.

## Talks and publications by the author (Апробация результатов исследования)

Published papers:

- 1. Турко Д. С. Феноменальный минимализм в онтологии самости // Антиномии, 2021, Т.21, №4, С. 7-30 [Tourko D. S. Phenomenal Minimalist Ontology of the Self. *Antinomii*, 2021, Vol. 21, №4, Р. 7-30].
- Турко Д. С. Опыт и его собственник: аргумент в пользу эгологизма // Вестник Московского университета. Серия 7: Философия, 2022, Т.2, С. 61-82 [Tourko D. S. Experience and its Owner: An Argument for Egologism. *The Moscow University Bulletin. Series 7. Philosophy.* 2022, Vol. 2, P. 61-82].
- Турко Д. С. Этика и проблема критериев субъектности // Вестник Томского государственного университета. Философия. Социология. Политология, 2022, Т.68 [Tourko D. S. Ethics and the problem of subjectivity criteria. *Tomsk State University Journal of Philosophy, Sociology and Political Science*. 2022, Vol. 68].
- Tourko D. We do not exist: The neurophilosophical stance against the 'self'.
   // Defining Nothingness. Conceptions of Negativity in Continental Philosophy. Edited by Tatiana Levina, Tatyana Lifintseva. Brill Publishing. 2024 (forthcoming).

Talks:

 Выступление на конференции 12th Salzburg Conference for Young Analytic Philosophy (SOPhiA 2022), Salzburg (online) 9 сентября 2022 г., название доклада: "Selfhood as an Emergent Property" [12th Salzburg Conference for Young Analytic Philosophy (SOPhiA 2022), Salzburg (online), talk: *Selfhood as an Emergent Property*, September 9, 2022].

- 2. Выступление на XI международной конференции Школы философии и культурологии НИУ ВШЭ «Способы мысли, пути говорения» 8-9 октября 2020, названия докладов: «Теория самости школы мадхьямика-шуньявада»; «Проблемы теории минимальной самости» [XI International Conference *The Modes of Thinking, The Ways of Speaking,* HSE School of Philosophy and Cultural Studies, talks: *The Madhyamika-Shunyavada School on the Self; Problems of the Minimal Self Theory,* October 8-9, 2020].
- 3. Выступление на круглом столе «Определяя ничто» Научноисследовательской группы «Метафизика и эпистемология» Школы философии и культурологии НИУ ВШЭ 7 декабря 2020, название доклада: «Пустое Я: есть ли референт у индекса первого лица?» [Defining Nothingness round table, Research Group Metaphysics and Epistemology, HSE School of Philosophy and Cultural Studies, talk: Empty Self: Does the first person indexical have a referent?, December 7, 2020].
- 4. Выступление на конференции «Актуальные проблемы аналитической философии 2021» (Томский государственный университет) 24 сентября 2021 г., название доклада: «Инфляционизм и дефляционизм в онтологии самости» [Actual Problems of Analytical Philosophy 2021 Conference, Tomsk State University (online), talk: Inflationism and deflationism in personal ontology, September 24, 2021].
- 5. Выступление на Двенадцатой международной конференции Школы философии и культурологии НИУ ВШЭ «Философия и культура в эпоху пандемии» 1 октября 2021 г., название доклада: «Датские норки, сентиентизм и моральная забота» [*Philosophy and Culture in Time of Pandemics*, Twelfth International Conference of the HSE School of Philosophy and Cultural Studies, talk: *Danish Minks, Sentientism and Moral Care,* October 1, 2021].

6. Конференция НИГа «Метафизика и эпистемология» Школы философии и культурологии ФГН НИУ ВШЭ «Defining Nothingness» 11 января 2022 г., название доклада: "We do not exist: The neurophilosophical stand against the 'self'" [Research Group *Metaphysics and Epistemology* Conference, HSE School of Philosophy and Cultural Studies, HSE School of Philosophy and Cultural Studies, talk: *We do not exist: The neurophilosophical stand against the 'self'*, January 11, 2022].

## DISSERTATION PRÉCIS (ОСНОВНОЕ СОДЕРЖАНИЕ ДИССЕРТАЦИИ)

The Introduction starts with stating the research field (personal ontology), two main research questions (the nature question and the reality question of the self or person), the central methodological problem (incompatibility of the three modes of description) and three reasons for the thesis' relevance (general interest in and importance of personal ontology today; practical relevance; and the necessity to clarify the relation between the three descriptive perspectives). The Literature review focuses on early modern classics (Descartes' Second Meditation, Locke's Essay II.27 and Hume's Treatise 1.4.6) and contemporary major works. It identifies several trends and camps in today's personal ontology and philosophical discourse on the self. Among the proponents of "brute-physical" views, animalists are the most prominent today. Animalist philosophers think each of us is numerically identical to a distinct human animal. Brain theorists identify ourselves with our brains or parts of brains. Some philosophers have an explicitly realist position about the self. They affirm the existence of the self as a mental entity distinct from animals and brains. Anti-realists deny the existence of such an entity. Among contemporary realist camps, we can count *soul-theorists*, *Neo-Cartesians* and *phenomenalists*. Soul-theorists and Neo-Cartesians see the self as an ontologically independent or really distinct entity; phenomenalists connect the reality of the self with the reality of consciousness and conscious experience. Antirealist theories are broken down into neurophilosophical anti-realism, Buddhist anti-realism; conventionalism, relativism and fictionalism among other positions. Neurophilosophical anti-realists either reduce the self to brain networks and capacities or deny its existence altogether utilizing neuroscientific and cognitivescientific arguments. Contemporary Buddhist anti-realists deny the reality of the self arguing from Buddhist philosophy. Derek Parfit's views and narrative theories of the self are also discussed. Narrativists see the self as something that comes into existence together with a unified autobiographical narrative. Finally, the review very briefly discusses other kinds of anti-realism as well as *Evolutionarybiological and neuroscientific theories of the self.* At the end of the review, the discussed positions are arranged into a chart to demonstrate a possible way of presenting the current field. Next, the *subject of inquiry, objectives, methodology, novelty,* and *principal positions* are offered, with detailed comments. The subject of this thesis' inquiry is the referent or referents of philosophical terms *self, person* and *subject.* The two primary investigated questions are, first, the reality of these terms' referent or referents (or: whether the terms are referring and not empty), and second, the nature of those. The Introduction concludes with an explanation of the thesis' *theoretical significance and possible practical applications* and *thesis structure*.

The first chapter *Experiential process-realism: statement and outline* serves, first, to introduce critical distinctions, terms and concepts; second, to properly approach the extremely vague and polysemic term *self*; third, to provide a brief outline and the initial argument for the position I champion – *experiential process-realism* (ExPR). I define ExPR as a *moderate* (non-substantivist), *naturalistic*, *egological intraneous*, *non-reductive*, *further-fact*, *emergentist*, *processual realism* about selves. ExPR is a synthetic account which provides answers to both the nature and reality question. I conclude the chapter with a review of the self's possible ontic categories – which double as possible ways to conceptualize the self.

The chapter kicks off with a discussion of terminology. The section *Person*, *ego*, *subject*, *I* and *self* explains how these family-related terms are typically used in contemporary philosophy. I use the terms *person* and *self* interchangeably, following the early modern example of Locke and Hume. The term *subject* is used specifically in the sense *subject of experience* and as such its referent will be distinct from those of *self* or *person*. It is also noted that the term *person* will be used strictly in the philosophical sense and not in the sense attached to this concept in psychology and other social sciences. The section *Humans vs persons* states that

it is commonplace in personal ontology literature to treat humans and persons (selves) as non-coextensive kinds. The reason for that is the modal distinctness of humans (human animals) and persons, meaning they are not associated necessarily. First, evidently there are humans who are not persons – possibly unborn and very small children and patients in a persistent vegetative state. it is very plausible there are non-human persons as well. The next section 'We': circularity or indefiniteness looks at how the term we is used in personal ontology literature. It criticizes Eric Olson's use of the term and concludes there is no way to define the term noncircularly. It is stated that instead of the question What are we? containing a vague term, we should ask more specific questions like What are we as animals? and What are we as persons essentially? The way we is used in the thesis implies we as persons unless stated otherwise. Next, in the section Intuitive definitions I consider numerous common definitions of the indeterminate term *self*. Many philosophers use the term *self* in a seemingly arbitrary manner, with an arbitrary set of properties. When starting philosophizing about the self, we should not impose some specific understanding of the term on the reader without a justification. To avoid that imposition and in order to establish an uncontroversial definition of the term *self*, I place some constraints on the candidates for definitions of selfhood: a definition must be intuitive and basic (not theory-laden). In order to supply such a definition, we need to run the candidate through a gauntlet of intuitive definitions; i.e. to check how the proposed understanding relates to our intuitive grasping of the term. I consider the following intuitive definitions: 1. Self is you and I and beings like us. 2. Self is identical to the human animal or organism. 3. It is the referent of "I". 4. Self must be able to think first-person thoughts. 5. Self is the subject of experience. 6. Self is the mental substance. 7. Self is the essence. 8. Self is the essential property. 9. Self is the essential property necessary for survival. Along the way, I make important distinctions between Substantivalism and nonsubstantivalism and Essentialist and pluralist conceptions. I also introduce the concept of Minimality or the minimal conception of the self. In the section Limitations of definitions I find that by itself none of the listed definitions is sufficient to provide a robust way for defining the term *self*. I suggest it can also be useful to start from square one and consider Selves as reflexive entities. I claim some things in the world have the property of *reflexivity*. By reflexivity I mean the property x of some system organized in a specific way such that x allows the system to represent its own internal states specifically for the system – for itself. "For" means a certain degree of integration and coding of this state in the internal organization of the system – such that the system achieves a degree of internal integration which translates into the state's epistemic isolation for an external observer. Reflexivity is a necessary (but probably not sufficient) condition for selfhood. Next, I discuss how reflexivity leads to a system's capacity for Spontaneous self-ascription of its own states. Finally, in the next section I arrive at the central claim of this thesis: *Selfhood is experientiality*. I briefly provide several reasons why this is a plausible way to understand selfhood, to be argued for in more detail later. In the next section Experientiality is sentience I identify these two properties: sentience in my definition means nothing but the capacity to have experience. I reply to an objection from circularity. The next section Experientiality and intuitive definitions makes good on the earlier promise and checks the proposed idea of selfhood against all intuitive definitions and ideas listed above. I show that experientiality fits these intuitions well, so there is no reason to reject the idea of selfhood as experientiality outright. In the section Substance, property, process? The problem of ontological placement I discuss the best way to conceive of the self's ontological category. Is our traditional understanding of selves as substances, or sharply bounded stable things, justified? I list four possible categories for the self: substance, organism, property and process. I suggest we should stick with the latter two because they have the most explanatory power and generate the least problems. I claim We are experiential processes because selves are just the processes of experiencing. In the final section The mongrel self I discuss four general distinct meanings or the term self, or as I call it selves considered at four distinct levels of ontological dependence: *creature-self, neural-self, phenomenal self* and *metaphysical self*. In the chapter summary, I provide what I call *The primitive argument* for the existence of selves which argues for the reality of selves from the reality of experience. The argument is basically the same as the Cartesian cogito.

Chapter two is the first chapter in the thesis' science-first block. In this chapter titled The question of nature: the evolved self I try to make a case for the reality of the self arguing on some evolutionary-biological and neurophilosophical (in short, science-first) grounds. The key thesis of this chapter is *naturalistic* realism: selves are best understood as naturally evolved entities. They represent a natural kind of things united by a property which is ascribed to them not conventionally, but due to their sharing the same kind. This is why selves are units of nature. In A true history of our selves I take up Daniel Dennett's statement that natural things, if they exist, ought to have a history of how they came to be. But the opposite is also true: if there is a true history of how something came to be, this thing exists (unless it had gone extinct earlier). I briefly relay Dennett's evolutionary account from his paper the Origins of Selves. According to Dennett, minimal selfhood was born when living organisms started to draw the line between themselves and the world. They did so as a means of distinguishing between the things more and less important for their own needs. Dennett concludes that the border which defines our selfhood exists for some organisms on multiple levels (e.g. on the level of its immune system or the social or tribal level). In the next section, I make an argument Against pragmatic eliminativism. The section's point is to refute the claims of philosophers like Metzinger and Olson who argue that the concept self can be discarded by science and philosophy because of its utter uselessness. I identify selves with experiencers, and experiencers with sentient organisms and conclude that to the extent a term singles out a part or aspect of reality, its utility is self-justifying. In that section, I make several important distinctions and clarifications to be used in The naturalistic argument for real selves. The argument is a long chain of definitions and identifications meant to show that if sentient organisms exist, the selves exist. Next, I consider what neuroscientist Antonio Damasio says on neural evolution of selves. I interpret Damasio's framework of the proto-self, the core self and extended consciousness as a strong support for the naturalistic thesis. The next section phenomenal dependence makes an important argument for the experientialist understanding of selves. Its point is that all self-capacities we typically ascribe to selves cannot manifest unless they are phenomenally realized – so they depend on the experiential capacity. This is why it is reasonable to consider experientiality as the base capacity and the essential property of selves. The naturalistic argument: second approach tries to reformulate the argument and clarify the meaning of the terms *natural* and *real*. It makes a crucial distinction between mental-as-content and mental-as-capacity. Drawing on that distinction, in the next section Hume and Kant on the self I interpret the famous part from Hume's Treatise 1.4.6.3 as the *Humean fallacy* – failure to distinguish between the two meanings of the mental (without claiming Hume made the mistake himself). Recasting the self from content to capacity or process is dubbed the Humean zoom out in this section. I reply to the objection about Hume's self-criticism in the Appendix to the *Treatise* and his failure to solve the problem of connexion for perceptions. Finally, I discuss how Kant solved the diachronic unity of consciousness problem via the concept of transcendental apperception. I do not think it is necessary to invoke the transcendental unity of apperception to account for the unity of consciousness. The naturalistic argument: third approach refines and summarizes what has been said about selves as units of nature so far. In the section Real, natural, artificial, conventional, fictional these different predicates are considered in relation to the self. I conclude that in the light of what has been said so far, selves are best thought of as real and natural and not artificial, conventional and fictional. In the final section Against conventional persons, I (somewhat drawing from Locke) argue against David Braddon-Mitchell and Kristie Miller's conventionalism about persons. I defend two primary claims: (1) because *person* is a mongrel concept, it can be analytically broken down into several kinds of person-like entities, each of which is individuated by distinct essential properties and persistence conditions; (2) for at least some of these kinds, their essential properties and persistence conditions are constituted by non-conventional facts and non-conventionally existing things. Therefore, a better way to understand personhood would be to adopt realistic pluralism about persons, as opposed to conventionalism.

Chapter three represents the second and final chapter in the naturalistic block of the thesis. It is entitled Being someone insubstantial: Against neurophilosophical anti-realism. It is a polemic against philosophers like Patricia Churchland, Daniel Dennett and especially Thomas Metzinger who argue against the existence of the self from neuroscientific and cognitive-scientific grounds. These philosophers either reduce the self to the brain's representational capacities (Churchland), eliminate it on illusionistic grounds (Metzinger) or recast it as an abstract useful fiction (Dennett). The first part of the chapter is a general introduction into the problem and relevant ideas. The section No-self from the neurophilosophical point of view describes the crux of the matter. *Neurophilosophical anti-realism: main points* lists views shared by neurophilosophical anti-realists. Problems discusses the cost of anti-realism: we will have to sacrifice a lot in our naive and philosophical understanding of persons if anti-realism is true. *No self in the brain?* sheds light on the important eponymous problem, also called the localization problem. The fact that there is no unified self to be found in the brain seems to shatter our naive presuppositions. I offer several rejoinders to the argument. I accuse neurophilosophical anti-realists of possibly ignoring phenomenological data, commiting a categorical mistake and phrenological thinking. In Alternatives, several non-reductionistic and noneliminativistic accounts of the self are considered, including those that argue from a naturalist perspective. The second part of the chapter (starting from the section Engaging Metzingerian eliminativism) deals specifically with Thomas Metzinger

and his book Being No One. The Metzingerian argument attempts to reconstruct the general logic of Metzinger's book and his phenomenal self-model theory (SMT). Each premise is commented on in a distinct subsection. The reconstructed argument goes like this: 1. We, human beings, are naturally predisposed to be naive Cartesian realists about the self (Naive Cartesians). 2. There is a true explanation for why we are naturally predisposed to be naive Cartesian realists about the self: Our phenomenal self-models (PSMs) are transparent and we confuse the content of our PSMs with our selves (Transparence: why we are naive realists). 3. There is a true causal explanation for the existence of PSMs: a PSM is a representational and teleofunctional tool naturally evolved by some biological organisms (Why PSMs exist). 4. (unstated premise: the debunking argument) If it is possible to explain why we have the idea of x without mentioning x, our ideas about x are not justified. 5. (from 1-4) It is possible to explain our idea of the self without mentioning the self (4-5. Metzinger's debunking argument). C1. Therefore, our ideas about the self are not justified. C2. Therefore, selves do not exist (anti-realism) and the concept self can be eliminated from science and philosophy (pragmatic eliminativism) (Cl-C2. Anti-realism and pragmatic eliminativism). The next section A Moorean argument? makes a case against Metzinger's view in the vein of Chalmers' Moorean argument for the existence of consciousness. It goes like this: 1. If eliminativism about the self is true, I do not exist. 2. I exist (claimed with certainty). C. Eliminativism about the self is false. Next, "We do not exist" discusses the conceptual kernel of Metzinger's position and concludes that if the proposed interpretation of his views is correct, his position is self-defeating and inconsistent. The section The phenomenal self and PSM asks the question whether Barry Dainton's conception of the self can be used to provide a complement or robust alternative to Metzinger's substantivalist picture of selves. A way to conjoin SMT and Dainton's PCS (potentially conscious self) is considered. In Selfprocesses: first approach, a first attempt to justify and describe selves as processes is made. Reductionism about processes is considered and rejected. I offer the first

working definition of the *processual self* (if x is identical to an experiential process at t, x is a self at t) and graft it onto Metzinger's SMT. Finally, I briefly consider the ontological distinction between "stable" things and processes and suggest it is ungrounded. In *Pragmatic justification: the self as a natural kind term*, I return to the idea of the pragmatic justification and consider whether the self can be construed as a natural-kind term drawing from the work of Nicholas Shea and Tim Bayne.

Chapter four The question of phenomenology: the self, the subject and its experience has to do with a phenomenological description of selfhood. Mostly it interprets and engages with the work of Galen Strawson. The central problem of this chapter is the reality of the subject of experience. I consider the subject to be a concept distinct from the self or person. Subject is understood to be a separate mental presence, a property of experience, a type of experiential content; or a watcher or owner, standing above and beyond its experience. A way to distinguish between the subject and the self/person I endorse is to assign them to different ontological categories: content and capacity/process, respectively. The first section Experience itself clarifies the meaning of the term experience. It approaches experience or phenomenal consciousness ostensively, pointing at its typically stated connected properties: what-it's-likeness, the first-person point of view, intentionality, IEM, representation, irreducibility, pre-reflective self-consciousness, mineness or for-me-ness, subjecthood and acquaintance. Property cluster for the natural kind 'self' somewhat repeats the test we subjected experientialism to in chapter one: it is investigated whether properties of experience just listed can also be considered properties of selves or subjects. Self-experience: what is it like to be someone? scrutinizes our experience of ourselves as subjects – as it is understood by Galen Strawson. Strawson thinks that in typical human self-experience we feel ourselves to be (1) a subject of experience (2) a thing of some sort which is (3) mental (4) synchronically and (5) diachronically single, is also (6) an agent (7) a character or personality and (8) distinct from the organism considered as a whole.

In addition to these properties, I list four additional properties of self-experience: presence, passing, subject-centeredness and detachment. Respectively, we feel we are present or existent as subjects; we experience the passing of the stream of consciousness; we perceive ourselves – our subjects of experience — as centered entities; and we are detached from our experiential streams. I also consider these properties natural and to a large extent universal in humans, although perhaps specific only for some subjects, i.e. possibly absent in non-human subjects. The next section *The elusive subject and modes of self-acquaintance* discusses a very widespread intuition that the subject cannot be directly given, found and grasped in experience by definition, just in virtue of being the subject and not the object of grasping. It discusses associated difficulties and possible solutions. It also discusses our ways of arguing for and against the existence of the subject. I conclude that although arguing for the existence of the subject from our selfexperience is legitimate, it leads to the subject being nothing more than a property of experience, never extending outside the duration of a single experiential episode. The second part of the chapter tries to prove exactly that. I claim the subject is best understood as a property of experience which is locked in a biconditional existential dependence with its experience, so that  $e \leftrightarrow s$ . The subject is an embedded real property of experience. To argue for that, I need to argue against any and all conceptions of the persistent detached subject, or the subjectas-watcher. The best way to do so is to argue that the subject of experience is The *transient subject*, meaning it does not last longer than its experience – the view I call transientism after Strawson's the transience view. I define Weak and strong transientism. I make The presentist argument for transientism. It goes like this: 1. All possible experience is limited to the present moment. (P1: Phenomenal Presentism). 2. No subject extends in time further than its experience. (P2: Thin Subjects). C. No subject extends in time further than the present moment. The argument argues that subjects of experience (The razor-thin subject and alternatives) are razor-thin subjects, meaning they exist only in the specious phenomenal present, as if on a razor edge. Next I discuss *Monistic and binary ontologies of consciousness*. The former make do without the subject in their understanding of consciousness, the latter posit it. I conclude only strong transientism satisfies the presentist constraint which requires that the subject does not extend beyond the present moment of its experience. Finally, I discuss *Some objections*: the *Endurantist objection* which tries to resolve the problem via a Lewis-style four-dimensional personal ontology and *the Continuous subject-time objection* which attacks the model of time-consciousness assumed in the chapter. *Transientist implications* discusses what it means for us if we are subjects of experience and strong transientism is true: we cannot ascribe to ourselves free will, moral responsibility, and other predicates characteristic of persons. I conclude that the subject of experience (content-level entity), but with the self or person instead (capacity-level).

Chapter five Shifting Persons, Real Persons: Buddhist Reductionism and Mereological Nihilism opens the metaphysical block of the thesis. Its main task is to engage with and refute a form of anti-realism extremely unlike neurophilosophical anti-realism: Buddhist Reductionism (BR) as it is reconstructed by Mark Siderits in his book Personal Identity and Buddhist Philosophy: Empty Persons (2017). The chapter begins with a discussion of the peculiarity of specifically Indian terms  $\bar{A}tman$  & pudgala. Buddhists and other Indian philosophers differentiated between the eternal, simple, passive, substantial and detached  $\bar{A}tman$  (self or soul) and the empirical pudgala (person), the bearer of psychological and possibly also bodily attributes. The central tenet of Buddhist philosophy is denying the existence of the soul-like self. At the same time, Buddhists are split on the reality of the person: there are eliminativists, realists (the pudgalavada school) and reductionists (the Abhidharma tradition). At the start, I attempt to reconstruct the Buddha's own views on the self, citing evidence from the Pali canon. I speculate that despite established Buddhist orthodoxy, it is perhaps not entirely clear what the Buddha's own views on the reality question were. In *Reductionism simpliciter*, I discuss what reductionism is generally. The first part of the chapter deals with BR arguments against the self. The Buddhist master *argument* against the reality of the self may be constructed like this: 1. There is no reason to believe the self exists. 1.1. There is no empirical reason (from phenomenological evidence). 1.2. There is no deduced reason (from inference). 2. Unless there is a reason to believe x exists, it is safe to conclude x does not exist. C. The self does not exist. I discuss *Objections to P1.1: the self can be experienced* at length and conclude that the persuasiveness of objections depends on how strongly we construe the subject of experience. The objections work as a way to argue for the existence of the content-level subject; but Buddhist anti-realist arguments work if we construe the self strongly, as a detached soul-like entity. Then I focus on *Objections to P1.2: the existence of the self can be inferred*. These are the following: a. The cogito. b. Self is the inter-personal individuator. c. Self is the intra-personal diachronic and synchronic unifier. d. Self is detached / self is the owner of its experience. Although I recognize some persuasiveness of these objections, in general I concur with BR that there is no good argument to posit the existence of an independent subject of experience or a fortiori a soul-like eternal entity which is the *Ātman*. However, the subject is perfectly real if meant to be just a property of experience. The second part of the chapter deals with Buddhist Reductionism about persons. Persons in Buddhism are understood very differently from selves. If selves are alleged soul-like entities, persons are understood as a causal series of psychophysical aggregates (skandhas). After a detailed discussion of BR arguments, I suggest that if skandha-series are in the constant state of flux, it is already sufficient for us to conceptualize selves as processes of some sort. In Self-processes: second approach (bureaucrat identity) I pitch a working definition of the process and set the process diachronic identity criterion. The two most important claims of BR about persons are the transience claim (persons do not persist) and the conventionality claim (persons are merely conventional entities

which do not exist at the ultimate level of reality). Both arguments in support of these claims are predicated on mereological nihilism (MN) - the view that composite things do not exist. So in order to defend the reality thesis from BR, I need to argue Against mereological nihilism. First I wonder Why would anyone believe that? As Siderits points out, BR motivation for MN (reducing persons and reducing suffering) is fundamentally different from that of contemporary Analytic proponents of MN (reaching "ontological seriousness", i.e. objective and subjective context-free description of reality modelled on the description in physics). I contemplate Arguments and objections for and against MN. My favoured way of arguing against MN is from *Emergent properties* of persons. Next, I claim that the proper way to individuate things in the world may be not based on their composition (as MN does), but on their functions (Functionality before *composition*). Functions of some wholes are irreducible to functions of their parts, so MN crumbles the moment we make this shift. I consider several Responses to the functionalist objection. Finally, having refuted MN, I have enough ground to argue Against transience and conventionality claims and with them - against Buddhist Reductionism about persons at large.

Chapter six *The question of metaphysics: the real self* is the second chapter in the metaphysical block. The purpose of this chapter is to discuss crucial metaphysical problems in relation to selves; and to clarify potential ExPR solutions to these problems. I begin by *Formulating realism* and by trying again to figure out *The meaning of 'real'*. I make some critical distinctions: *1. Strong vs moderate realism* which develops the substantivalist-non-substantivalist binary from ch. 1 and also discusses Parfit's Reductionism about persons and the ambiguity of terms like *separately existing* and *distinct*; *2. Egologism and non-egologism* has to do with the subject's presence in consciousness; *3. Reductive vs non-reductive theories* returns to Parfit's Reductionism and promises to refute it in ch. 7. Finally, *4. Self-experience (sense of self) vs ontic self* develops the distinction and relation between the phenomenally present self and independently existing self discussed in two previous chapters. In Self-processes: third approach (fire and fuel), I return to the topic broached in the previous chapter: the nature of processes and selves as processes. I take up the Buddhist metaphor of fire and fuel to illustrate how it all might work. I define the three kinds of self-processes depending on their ontic category: creature, neural stuff and phenomenal stuff. In his Essay, Locke stipulated different conditions of diachronic identity for different kinds of things (material constitution plus mereological essentialism for inanimate things, sameness of life for organisms, and sameness of consciousness for persons). I accept this view and posit an additional continuity constraint for all living and conscious processes: preserving an organism's or thinking thing's identity and survival requires for the process to be continuous and uninterrupted. In *The bridge* problem: every night I die, I engage with one of the principal problems for all experientialist accounts of personal ontology: the problem of accounting for periods of unconsciousness. I briefly consider Dainton's solution (he recast the essentially conscious self as the potentially conscious self) and find it problematic. My solution has to do with analytically separating process-interruption from process-suspension. A self-process is suspended if (1) it stops realizing its relevant property and (2) its direct lower-order base process *continues* to realize its relevant property. So if there is an effective low-order base realizer to support the essential higher-order property, the process associated with the essential property is not interrupted, but merely suspended. In *Waterfall-self*, I play with a water metaphor for the experiential process to pair it up with the original Buddihst fire-fuel metaphor. In The processual cogito and fission, I return to the classical Cartesian argument and consider it from an ExPR point of view. I rewrite it to make it consist of just one premise: 1. There is experiencing going on. C. Experiencing exists. I am identical to that experiencing. I consider how ExPR can be applied to solving the problem of fission cases for diachronic personal identity. In the end, in line with transientism, I conclude that each of our occurrent stages is not numerically identical to any of our past or future stages. The final part of this chapter returns

directly to the main question of personal ontology: What are we in the end? First, I ask the question Are we brains (or brain parts)? While in some sense it is perfectly legitimate to conceive ourselves as neural selves, there are important objections which ultimately force us to answer the question in the negative. Next, I try to resolve the personalist-animalist debate: the contention whether we are first and foremost animals or persons (Our kind: animals or persons?). Eventually I conclude that there is no factual answer to the question. When choosing between Animal-priority vs person-priority, all that can force our hand is our own valueladen considerations. Persons are prior for us and we are primarily persons only in virtue of persons being value-laden beings or value-sensitive agents. Because we prioritize some values, we also prioritize persons. As a final farewell for animalism, I consider its trademark The thinking animal argument. Its point is to show that we as thinkers are identical to animals because otherwise it would lead to overcrowding of the same point in space with too many thinkers. I attempt to apply ExPR in trying to attack the argument. My rewrite of it suggests animals are not their thinking; animals generate the process of thinking and we are identical to that process. So we are not identical to animals, strictly speaking. The section *Back* to the terms discusses the utility of using indeterminate concepts like self, person, ego and subject and interprets Shoemaker's famous brain transplant thought experiment without resorting to those indeterminate terms. At the end, I return to the question of not what we are simpliciter, but what we are essentially as persons. I offer The survivability test as a means to check our intuitions about our fundamental nature as persons. Following Dainton, I offer three speculative scenarios which imply we are intuitively inclined to see ourselves as conscious entities and not animals or psychological content-bearers. Finally, I make The value argument for experientialism which is distinct from but somewhat conceptually connected to the earlier phenomenal dependence argument. The value argument says even if there is no metaphysical fact of the matter concerning what we are as persons (just like, as I believe, there is no factual answer to the question *What are we*? simpliciter), because value depends on experientiality, it is meaningful to single out experientiality as the essential property of persons, even if persons are entirely conventional entities. So here we can make a deontic argument: we are value-sensitive beings and we cherish our ability to be sensitive to value; so preserving the one aspect of our nature which makes value meaningful is our duty before ourselves.

The final chapter seven of the thesis, Irreducible persons, is the most adventurous one. It covers a controversial thesis that persons are irreducible and some facts about persons are further-facts. In the Salience of I\*, I define I\* as the occurrent subject of experience; the one that stands in a relation of direct selfacquaintance. Next, I try to distinguish between *Persons vs I\**. It seems there is just one I\* in existence if we suspend our belief in other minds. There are two ways to define I\*: either descriptively (1a. Direct self-acquaintance and being occurrently conscious) or ostensively, via a demonstrative indexical (2a. Being this subject). Twin-I offers a scenario to illustrate this idea. Suppose there is Twin-I living on Twin-Earth. Suppose also Twin-Earth and all of its inhabitants are physically and phenomenally qualitatively identical to Earth and all its inhabitants. I\* and Twin-I are different persons, so we are numerically distinct. But it is also impossible for us to be really qualitatively identical: I\* am the I\*-property bearer and Twin-I is not. If I\* is instantaneously destroyed and substituted with a clone of Twin-I, this event will have no effect on anything on Earth. Earth and Twin-Earth will continue to be qualitatively identical. But such a substitution would obviously mean the world for I\*, as I\* would be obliterated. It follows that qualitatively identical persons are inconceivable, given as long as one of the relata in the twoperson identity relation is the I\*-bearer. The paradox of many I\*'s is a paradox that all these three incompatible propositions seem true: (a). I\* is a singular term. (b). If x is a person, x is an I\*-bearer, at least for itself. (c). There is more than one person. All three statements seem to be very plausible by themselves, but it is impossible to hold all three to be true simultaneously. a & b is solipsism. b & c is common sense. a & c implies persons do not necessarily have centered perspectives, so a & c is a situation of one I\* and decentered, perspectiveless zombies. Finally, I make an involved *knowledge argument for irreducible persons*. After multiple refinements and restatements, the argument concludes that being I\* is a further fact. The justification for this is that an omniscient observer (we can call it the Laplacian Demon) who could know any fact it wishes to know could not possible know some facts containing I\* (e.g. which one of the multiple persons in existence is I\*). I consider several *Responses and objections* to the argument, including solipsism, us being Leibnizian monads and the obvious objection that the argument messes up the logic of the first-person indexical. I reject this objection and endorse the emergent-property explanation of persons being irreducible.

In the *Conclusion*, I consider *Some implications* of the thesis' theoretical results. These results are relevant for topics like *Death, immortality and transhumanism, Artificial persons, Moral status of persons* and *Moral status of non-human animals*. The latter I discuss in some detail, considering different ways to construe the criteria of our moral concern, including the criteria based on similarity and subjectivity type. I endorse *sentientism* – the view that moral worth is predicated on sentience. Therefore, if experientialism is true, any person and any conscious animal, no matter how significant or neurally primitive, has moral worth. I conclude the thesis with the idea that if it is true that we are essentially experiencers, it means consciousness constitutes our lifeworld as persons. An individual person. Consciousness is where our life begins, goes and ends. Consciousness is to be cherished.

## REFERENCES

- 1. Albahari, M. (2006). *Analytical Buddhism*. Palgrave Macmillan UK. https://doi.org/10.1057/9780230800540
- Albahari, M. (2010). Nirvana and Ownerless Consciousness. In Self, No Self? Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199593804.003.0004
- Albahari, M. (2014). Insight Knowledge of No Self in Buddhism: An Epistemic Analysis. *Philosopher's Imprint*, 14(21). http://hdl.handle.net/2027/spo.3521354.0014.021
- Bailey, A. M. (2015). Animalism. *Philosophy Compass*, 10(12), 867–883. https://doi.org/10.1111/phc3.12263
- Bailey, A. M. (2016). You Are an Animal. *Res Philosophica*, *93*(1), 205–218. https://doi.org/10.11612/resphil.2016.93.1.9
- Bailey, A. M. (2017). Our animal interests. *Philosophical Studies*, 174(9), 2315–2328. https://doi.org/10.1007/s11098-016-0800-6
- Bailey, A. M., & Elswyk, P. van. (2021). Generic Animalism. *Journal of Philosophy*, *118*(8), 405–429. https://doi.org/10.5840/jphil2021118829
- Bailey, A. M., Thornton, A. K., & van Elswyk, P. (2021). Why animalism matters. *Philosophical Studies*, *178*(9), 2929–2942. https://doi.org/10.1007/s11098-020-01593-x
- Bihan, B. L. (2019). The No-Self View and the Meaning of Life. *Philosophy East and West*, 69(2), 419–438. https://doi.org/10.1353/pew.2019.0033
- 10.Blatti, S. (2012). A New Argument for Animalism. *Analysis*, 72(4), 685–690. https://doi.org/10.1093/analys/ans102
- 11.Blatti, S. (2014). Animalism. In E. Zalta (Ed.), *Stanford Encyclopedia of Philosophy*.

- 12.Blatti, S., & Snowdon, P. F. (2016). *Animalism: New Essays on Persons, Animals, and Identity*. Oxford University Press UK.
- 13.Braddon-Mitchell, D., & Miller, K. (2004). How to Be a Conventional Person. *Monist*, 87(4), 457–474. https://doi.org/10.5840/monist200487423
- 14.Braddon-Mitchell, D., & Miller, K. (2020). Surviving, to Some Degree. *Philosophical Studies*, 177(12), 3805–3831. https://doi.org/10.1007/s11098-019-01410-0
- 15.Carr, D. (1991). *Time, narrative, and history* (1. Midland book ed). Indiana Univ. Press.
- 16.Chadha, M. (2017). No-Self and the phenomenology of agency. *Phenomenology and the Cognitive Sciences*, 16(2), 187–205. https://doi.org/10.1007/s11097-016-9455-1
- 17.Chadha, M. (2018). No-Self and the Phenomenology of Ownership. *Australasian Journal of Philosophy*, 96(1), 14–27. https://doi.org/10.1080/00048402.2017.1307236
- 18.Chadha, M. (2019). Reconstructing memories, deconstructing the self. *Mind & Language*, 34(1), 121–138. https://doi.org/10.1111/mila.12204
- 19.Chadha, M. (2021). Eliminating Selves and Persons. Journal of the American Philosophical Association, 7(3), 273–294. https://doi.org/10.1017/apa.2020.27
- 20.Chadha, M., & Nichols, S. (2022). Self-Control without a Self. *Australasian Journal* of *Philosophy*, 1–18. https://doi.org/10.1080/00048402.2022.2049833
- 21.Christman, J. (2004). Narrative Unity as a Condition of Personhood. *Metaphilosophy*, 35(5), 695–713. https://doi.org/10.1111/j.1467-9973.2004.00345.x
- 22.Churchland, P. (2013). *Touching a Nerve: Exploring The Implications Of The Self As Brain*. WW Norton.

- 23.Churchland, P. S. (2002a). *Brain-Wise: Studies in Neurophilosophy* (Illustrated edition). Bradford Books.
- 24.Churchland, P. S. (2002b). Self-Representation in Nervous Systems. *Science*, 296(5566), 308–310. https://doi.org/10.1126/science.1070564
- 25.Churchland, P. S. (2011). The Brain and Its Self. Proceedings of the American Philosophical Society, 155(1), 41–50.
- 26.Corcoran, K. (Ed.). (2001). Soul, body, and survival: Essays on the metaphysics of human persons. Cornell University Press.
- 27.Dainton, B. (2008). *The Phenomenal Self*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199288847.001.0001
- 28.Dainton, B. (2012). Selfhood and the Flow of Experience. Grazer Philosophische Studien, 84(1), 161–199. https://doi.org/10.1163/9789401207904 009
- 29.Dainton, B., & Bayne, T. (2005). Consciousness as a guide to personal persistence. Australasian Journal of Philosophy, 83(4), 549–571. https://doi.org/10.1080/00048400500338856
- 30.Davenport, J. J. (2012). Narrative Identity, Autonomy, and Mortality: From Frankfurt and MacIntyre to Kierkegaard. Routledge.
- 31.Deikman, A. (1996). 'I' = Awareness. Journal of Consciousness Studies, 3(4), 350–356.
- 32.Dennett, D. (1992). The self as a center of narrative gravity. In F. Kessel & D. Cole (Eds.), *Self and consciousness: Multiple perspectives* (pp. 103–115). Erlbaum.
- 33.Dennett, D. C. (1989). The Origins of Selves. Cogito, 3(3), 163–173. https://doi.org/10.5840/cogito19893348
- 34.Dennett, D. C. (1991). *Consciousness explained* (1st ed). Little, Brown and Co.

- 35.Dreyfus, G. (2010). Self and Subjectivity: A Middle Way Approach. In *Self, No Self?* Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199593804.003.0005
- 36.Duncan, M. (2021). Animalism is Either False of Uninteresting (Perhaps Both). *American Philosophical Quarterly*, 58(2), 187–200. https://doi.org/10.2307/48614005
- 37.Foster, J. (1991). The Immaterial Self: A Defence of the Cartesian Dualist Conception of the Mind. Routledge.
- 38.Gallagher, S., & Zahavi, D. (2020). The Phenomenological Mind. Routledge.
- 39.Geddes, A. (2013). Think Twice, It's All Right: Animalism, Disunity and The Self. *Proceedings of the Aristotelian Society*, *113*(3pt3), 371–380. https://doi.org/10.1111/j.1467-9264.2013.00359.x
- 40.Gillihan, S. J., & Farah, M. J. (2005). Is Self Special? A Critical Review of Evidence From Experimental Psychology and Cognitive Neuroscience. *Psychological Bulletin*, 131(1), 76–97. https://doi.org/10.1037/0033-2909.131.1.76
- 41.Hoeltje, M., Schnieder, B., & Steinberg, A. (Eds.). (2013). Varieties of Dependence: Ontological Dependence, Grounding, Supervenience, ResponseDependence. Philosophia Verlag GmbH. https://doi.org/10.2307/j.ctv2nrzhj9
- 42.Hood, B. (2012). The Self Illusion: Why There is No 'You' Inside Your He.... Constable & Robinson. https://www.goodreads.com/book/show/13384559the-self-illusion
- 43.Hudson, H. (2001). A Materialist Metaphysics of the Human Person. Cornell University Press. https://doi.org/10.7591/9781501725715
- 44.Hudson, H. (2007). I Am Not an Animal! In P. Van Inwagen & D. Zimmerman (Eds.), *Persons: Human and Divine* (pp. 216–234). Oxford University Press. https://philpapers.org/rec/HUDIAN

- 45.Hume, D. (1978). A Treatise of Human Nature [1978 reissue] (L. A. Selby-Bigge, Ed.; 2. ed., repr). Oxford University Press (Clarendon Press). https://doi.org/10.1093/actrade/9780198245872.book.1
- 46.Inwagen, P. van, & Zimmerman, D. (2007). *Persons: Human and Divine*. Clarendon Press.
- 47.Jennings, C. D. (2020). Practical Realism About the Self. In L. R. G. Oliveira & K. Corcoran (Eds.), *Common Sense Metaphysics: Essays in Honor of Lynne Rudder Baker*. Routledge.
- 48.Johnston, M. (2007). Human beings revisited: My body is not an animal. InD. Zimmerman (Ed.), *Oxford Studies in Metaphysics* (Vol. 3, pp. 33–74).
- 49.Johnston, M. (2016). Remnant Persons: Animalism's Undoing. In S. Blatti & P. F. Snowdon (Eds.), *Animalism: New Essays on Persons, Animals, and Identity* (pp. 89–127). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199608751.003.0005
- 50.Kenny, A. (1988). The Self.
- 51.Krueger, J. W. (2010). The Who and the How of Experience. In Self, NoSelf?OxfordUniversityPress.https://doi.org/10.1093/acprof:oso/9780199593804.003.0002
- 52.Lewis, D. (1983). Survival and Identity. In D. Lewis, *Philosophical Papers Volume I* (1st ed., pp. 55–77). Oxford University PressNew York. https://doi.org/10.1093/0195032047.003.0005
- 53.Locke, J. (1975). The Clarendon Edition of the Works of John Locke: An Essay Concerning Human Understanding (P. H. Nidditch, Ed.). Oxford University Press (Clarendon Press). https://doi.org/10.1093/actrade/9780198243861.book.1
- 54.Lowe, E. J. (1991). Substance and Selfhood. *Philosophy*, 66(255), 81–99.
- 55.Lund, D. H. (1994). *Perception, Mind and Personal Identity* (1st edition). University Press Of America.

- 56.Lund, D. H. (2005). *The conscious self: The immaterial center of subjective states*. Humanity Books.
- 57.Lund, D. H. (2009). Persons, Souls and Death: A Philosophical Investigation of an Afterlife. McFarland & Company.
- 58.MacIntyre, A. C. (2007). *After virtue: A study in moral theory* (3rd ed). University of Notre Dame Press.
- 59.MacKenzie, M. (2010). Enacting the Self: Buddhist and Enactivist Approaches to the Emergence of the Self. In *Self, No Self?* Oxford University Press.

https://doi.org/10.1093/acprof:oso/9780199593804.003.0010

- 60.MacKenzie, M. (2022). Luminous Cognizance: Toward a Buddhist Model of Consciousness.
  https://www.academia.edu/27785906/Luminous\_Cognizance\_Toward\_a\_Bu ddhist\_Model\_of\_Consciousness?
  auto=download&email work card=download-paper
- 61.MacKenzie, M. D. (2007). The Illumination of Consciousness: Approaches to Self-Awareness in the Indian and Western Traditions. *Philosophy East and West*, 57(1), 40–62. https://doi.org/10.1353/pew.2007.0006
- 62.Martin, R., & Barresi, J. (2006). *The Rise and Fall of Soul and Self: An Intellectual History of Personal Identity*. Columbia University Press.
- 63.McClelland, T. (2019). Against Virtual Selves. *Erkenntnis*, 84(1), 21–40. https://doi.org/10.1007/s10670-017-9945-8
- 64.McGinn, C. (1996). *The character of mind: An introduction to the philosophy of mind* (2nd ed.). Oxford University Press.
- 65.Merricks, T. (2001). Realism About Personal Identity Over Time. *Noûs*, 35(s15), 173–187. https://doi.org/10.1111/0029-4624.35.s15.9
- 66.Metzinger, T. (2003). Being no one: The self-model theory of subjectivity. MIT Press.

- 67.Metzinger, T. (2007). Empirical perspectives from the self-model theory of subjectivity: A brief summary with examples. In *Progress in Brain Research* (Vol. 168, pp. 215–278). Elsevier. https://doi.org/10.1016/S0079-6123(07)68018-2
- 68.Metzinger, T. (2010a). *The Ego Tunnel: The Science of the Mind and the Myth of the Self* (Reprint edition). Basic Books.
- 69.Metzinger, T. (2010b). The Self-Model Theory of Subjectivity: A Brief Summary with Examples. *HUMANA*. *MENTE Journal of Philosophical Studies*, 14, 25–53.
- 70.Metzinger, T. (2011). The no-Self alternative. In S. Gallagher (Ed.), *The Oxford Handbook of the Self* (pp. 279–296). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199548019.003.0007
- 71.Moreland, J. P. (2014). *The Soul: How We Know It's Real and Why It Matters* (1st edition). Moody Publishers.
- 72.Moreland, J. P., & Rae, S. B. (2000). *Body & Soul: Human Nature the Crisis in Ethics* (3rd Paperback Edition). IVP Academic.
- 73.Nagel, T. (1971). Brain bisection and the unity of consciousness. *Synthese*, *22*(3–4), 396–413. https://doi.org/10.1007/BF00413435
- 74.Olson, E. T. (1999). *The Human Animal: Personal Identity without Psychology* (Revised ed. edition). Oxford University Press.
- 75.Olson, E. T. (2003). An Argument for Animalism. In R. Martin & J. Barresi (Eds.), *Personal identity*. Blackwell.
- 76.Olson, E. T. (2007). What Are We?: A Study in Personal Ontology. New York: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780195176421.001.0001
- 77.Olson, E. T. (2015). On Parfit's View That We Are Not Human Beings. *Royal Institute of Philosophy Supplement*, 76, 39–56.
  https://doi.org/10.1017/S1358246115000107

- 78.Parfit, D. (1971). Personal Identity. The Philosophical Review, 80(1), 3. https://doi.org/10.2307/2184309
- 79. Parfit, D. (1984). Reasons and Persons. Clarendon Press.
- 80.Parfit, D. (2012). We Are Not Human Beings. *Philosophy*, 87(1), 5–28. https://doi.org/10.1017/S0031819111000520
- 81.Puccetti, R. (1973). Brain Bisection and Personal Identity. *The British Journal for the Philosophy of Science*, 24(4), 339–355. https://doi.org/10.1093/bjps/24.4.339
- 82. Ricoeur, P. (1994). Oneself as Another. University of Chicago Press.
- 83.Rigterink, R. J. (1980). Puccetti and Brain Bisection: An Attempt at Mental Division. *Canadian Journal of Philosophy*, 10(3), 429–452. https://doi.org/10.1080/00455091.1980.10715735
- 84.Rudd, A. (2009). In Defence of Narrative. *European Journal of Philosophy*, *17*(1), 60–75. https://doi.org/10.1111/j.1468-0378.2007.00272.x
- 85.Rudd, A. (2012). *Self, Value, and Narrative: A Kierkegaardian Approach*. OUP Oxford.
- 86.Schechtman, M. (1996). *The Constitution of Selves*. Cornell University Press.
- 87.Schechtman, M. (2007). Stories, Lives, and Basic Survival: A Refinement and Defense of the Narrative View. *Royal Institute of Philosophy Supplement*, 60, 155–178. https://doi.org/10.1017/S1358246107000082
- 88.Schechtman, M. (2014). *Staying Alive: Personal Identity, Practical Concerns, and the Unity of a Life.* OUP Oxford.
- 89.Shoemaker, D. (2016). The Stony Metaphysical Heart of Animalism. In S. Blatti & P. F. Snowdon (Eds.), *Animalism* (pp. 303–328). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199608751.003.0015
- 90.Shoemaker, S. (1963). Self-knowledge and self-identity. Cornell University Press,.
- 91. Shoemaker, S., & Swinburne, R. (1984). Personal Identity. Blackwell.

- 92.Siderits, M. (2010). Buddhas as Zombies:A Buddhist Reduction of Subjectivity. In Self, No Self? Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199593804.003.0012
- 93.Siderits, M. (2011). Buddhist Non-Self: The No-Owner's Manual. In S. Gallagher (Ed.), *The Oxford Handbook of the Self* (pp. 297–315). Oxford University

https://doi.org/10.1093/oxfordhb/9780199548019.001.0001

- 94.Siderits, M. (2017). Personal Identity and Buddhist Philosophy: Empty Persons (1st ed.). Routledge. https://doi.org/10.4324/9781315247052
- 95.Siderits, M. (2019). Persons and Selves in Buddhist Philosophy. In M. Siderits, *Persons* (pp. 301–325). Oxford University Press. https://doi.org/10.1093/oso/9780190634384.003.0012
- 96.Siderits, M., Thompson, E., & Zahavi, D. (Eds.). (2010). Self, No Self?: Perspectives from Analytical, Phenomenological, and Indian Traditions. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199593804.001.0001
- 97.Snowdon, P. (2009). The Self and Personal Identity. In J. Shand (Ed.), *Central Issues of Philosophy* (pp. 121–137). Wiley-Blackwell.
- 98.Snowdon, P. F. (1995). Persons, Animals and Bodies. In J. L. Bermudez, A. J. Marcel, & N. M. Eilan (Eds.), *The Body and the Self*. MIT Press.
- 99.Snowdon, P. F. (2004). Persons, Animals, and Ourselves. In T. Crane & K. Farkas (Eds.), *Metaphysics: A Guide and Anthology*. Oxford University Press.
- Snowdon, P. F. (2014). Persons, Animals, Ourselves. Oxford University Press.
   https://doi.org/10.1093/acprof:oso/9780198719618.001.0001
- 101. Strawson, G. (1997). 'The Self'. Journal of Consciousness Studies, 4(5–6), 405–428.

- 102.Strawson, G. (2009). Selves: An Essay in Revisionary Metaphysics.OxfordUniversityhttps://doi.org/10.1093/acprof:oso/9780198250067.001.0001
- 103. Strawson, G. (2011). The Minimal Subject. In S. Gallagher (Ed.), *The Oxford Handbook of the Self* (pp. 253–278). Oxford University Press.
- 104. Strawson, G. (2017). *The Subject of Experience*. Oxford University Press.
- 105. Swinburne, R. (1986). *The Evolution of the Soul*. Oxford University Press.
- 106. Swinburne, R. (2013). *Mind, Brain, and Free Will* (1st edition). Oxford University Press.
- 107. Swinburne, R. (2019). Are We Bodies or Souls? Oxford University Press.
- 108. Taylor, C. (1989). Sources of the self: The making of the modern *identity*. Harvard University Press.
- 109. Tekin, Ş. (2015). Self, Philosophical Considerations. In R. L. Cautin & S. O. Lilienfeld (Eds.), *The Encyclopedia of Clinical Psychology* (pp. 1–6). John Wiley & Sons, Inc. https://doi.org/10.1002/9781118625392.wbecp519
- 110. Thomson, J. J. (1997). People and Their Bodies. In T. Sider, J. Hawthorne, & D. W. Zimmerman (Eds.), *Contemporary Debates in Metaphysics*. Blackwell.
- 111. Toner, P. (2011). Hylemorphic animalism. *Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition*, 155(1), 65–81. https://doi.org/10.1007/s11098-010-9522-3
- 112. Unger, P. (2000). The Survival of the Sentient. *Philosophical Perspectives*, 14, 325–348.
- 113. Vogeley, K., & Gallagher, S. (2011). Self in the brain. In *The Oxford Handbook of the Self* (pp. 111–136). Oxford University Press.

- 114. Weir, R. S. (2023). *The Mind-Body Problem and Metaphysics: An Argument from Consciousness to Mental Substance*. Taylor & Francis.
- Wiggins, D. (2012). Identity, Individuation and Substance. *European Journal of Philosophy*, 20(1), 1–25. https://doi.org/10.1111/j.1468-0378.2012.00516.x
- 117. Гаспаров, И. Г. (2021). Как душа могла бы быть формой тела? Аристотелевско-схоластический подход к вопросу о метафизической природе человеческой личности / How soul could be the form of body? An Aristotelian-Scholastic approach to the question of the metaphysical nature of the human person. *Philosophy Journal*, 65–81. https://doi.org/10.21146/2072-0726-2021-14-4-65-81
- Нехаев, А. В. (2021). Почему так важно быть мёртвым? Омский Научный Вестник. Серия «Общество. История. Современность», 6(3), 90–107. https://doi.org/10.25206/2542-0488-2021-6-3-90-107