Introduction

JeeLoo Liu

In our everyday activities, the self is ever-present in the back of our minds. We remember what we did the moment before and we think about what we want to do next; we feel happy and energetic, or bored and tired; we have a sense of our goals when we act; we think about what we would like to eat for dinner and we know what our favorite TV show is. In our interactions with others, we think about how they see us, whether they like us or are impressed by us. We have certain emotions related to this keen awareness of ourselves: we feel embarrassed, remorseful, ashamed, proud, or confident because of things we have done or did not do. We see ourselves as continuous in time: what happened to us in the past affects who we are and what we believe now; we make plans for the future because we believe that the future self will be us and will be affected by our current plan and behavior. Even though we do not have an internal mirror to see ourselves, our every thought seems to revolve around the sense of a self. But what is the self? How is our sense of the self established in the first place?

The title of this book is *Consciousness and the Self*. The main focus of the collected essays is not to establish a metaphysical claim about the existence or the nature of the self, but to investigate the connection between our conscious life and our sense of the self; in other words, the *phenomenological* routes to the self. Whether or not we can establish the existence of a self, we undeniably have a sense of our self in our daily conscious life, in our reflections, sensations, discourses, memories, and our life plans. Phenomenally, I know *what it is like to be me*, and no one else can have my phenomenal awareness of my self. My self and my awareness of myself seem essentially intertwined.

Descartes' famous *cogito, ergo sum* points out the necessary presence of a self in consciousness: The *I* necessarily exists as the subject of thinking. It is a thinking thing embodied in the act of thinking itself. Descartes says, "So after considering everything very thoroughly, I must finally conclude that this proposition, I am, I exist, is necessarily true whenever it is put forward

by me or conceived in my mind" (Descartes [1641] 1984, 17). The *I* is essential to consciousness; all thinking requires a thinker. If I reflect on my act of thinking, then I know that there must be an *I* doing the thinking. Descartes' argument can be interpreted as deriving certainty of the second-order self-reflective thought on the first-order thinking. The certainty is only established for *synchronic* unity in thinking – at each single moment of thinking, I know that I exist as the thinker.

John Locke, however, emphasizes the *diachronic* unity of the self in all our conscious moments:

> When we see, hear, smell, taste, feel, meditate, or will anything, we know that we do so. Thus it is always as to our present sensations and perceptions: and by this every one is to himself that which he calls self: – it not being considered, in this case, whether the same self be continued in the same or diverse substances. For, since consciousness always accompanies thinking, and it is that which makes every one to be what he calls self, and thereby distinguishes himself from all other thinking things, in this alone consists personal identity, i.e. the sameness of a rational being, and as far as this consciousness can be extended backwards to any past action or thought, so far reaches the identity of that person; it is the same self now it was then; and it is by the same self with this present one that now reflects on it, that that action was done. (Locke [1689] 1975, 335)

In this passage, Locke sums up the connection between consciousness and the self: *consciousness is that which makes up the self* – both the presence of the self in occurrent conscious moments, and the persistence of the self in consciousness extended backwards in time. Self-consciousness is essential to personhood, and personal identity is grounded on one's memories or ownership of past deeds: *I was the one who did this*. The ascription of a past life to oneself is based on the assumption of a self that persists from the past to the present. Such persistence of the self, according to Locke, is sustained in consciousness alone, wherever that consciousness resides – whether in a single immaterial substance, or a succession of immaterial substances, or even, as he was willing to consider, in the brain, a succession of complexes of material substances.

Nevertheless, David Hume raises skepticism about our ability to perceive the self. Hume says,

> All [our particular perceptions] are different, and distinguishable, and separable from each other, and may be separately considered, and may exist separately, and have no need of any thing to support their existence. After what manner therefore do they belong to self, and how are they connected with it? For my part, when I enter most intimately into what I call myself, I always stumble on some particular perception or other, of heat or cold, light or shade, love or hatred, pain or pleasure.

I never can catch myself at any time without a perception, and never can observe any thing but the perception. (Hume [1739] 2000, 252)

Hume's skepticism is about our supposedly intimate consciousness of what we call "the self." He holds that for any real idea, there must be a preceding impression from which one derives the idea. However, our impressions are of our constantly changing perceptions, sensations, passions and emotions, and the like. We do not have such an impression of an everlasting self, persisting through all these changing impressions. If we do not have an impression of the self, then we cannot be said to have a clear idea of the self. The mind is like a kind of theater, according to Hume, "where several perceptions successively make their appearance; pass, repass, glide away, and mingle in an infinite variety of postures and situations" (Hume [1739] 2000, 253). There is nothing unified, invariable and persistent behind all these perceptions that we can call "the self."

Hume's comment points out a paradoxical double role that the self plays in self-awareness: the self as the *subject* and the self as the *object*. According to Ludwig Wittgenstein, there are two different uses of the word 'I': "the use as object" and "the use as subject." When the word 'I' is used to pick out a particular person of which a description can be judged to be true or false, as in "I have a bump on my forehead," it is used as object; on the other hand, when 'I' is used to report a sense of agency, as in "I think it will rain," it is used as subject (Wittgenstein [1958] 1969, 66–67). In self-awareness or self-knowledge, both uses seem to be present. "*I* believe that *I* am the tallest person in the class"; "*I* know that *I* am not sad about her departure." How can there be two selves indicated in these self-reports, or is it just one self who knows, perceives, thinks about, or is aware of, *the same self*? How can the same self be both the knower and the known?

Hume's claim can be taken to be a rejection of the self as an object of knowledge, or a rejection of any such unified entity as *the self*. Both rejections have their defenders. In *Notebooks*, Wittgenstein exclaims: "The I is not an object. I objectively confront every object. But not the I" (Wittgenstein 1984, 80e). In *Beyond Good and Evil*, Nietzsche questions the Cartesian certainty in deducing the existence of an *I* from the act of thinking: "The philosopher must say to himself, 'When I dismantle the process which is expressed in the sentence "I think," I come upon a series of daring assertions whose grounding is difficult, perhaps impossible – for example, that *I* am the one who thinks, that there must be some general something that thinks, that thinking is an action and effect of a being which is to be thought of as a cause, that there is an "I."' . . ." (Nietzsche [1886]

2009, Part 1, 16). Nietzsche also questions the unity of consciousness: "we always have only a semblance of unity" (Nietzsche [1901] 1968, 489). His hypothesis of the self is "the subject as multiplicity": the subject *I* is merely the sum of multiple perspectives, interpretations, and drives. Our customary use of the first-person word 'I' may have begun as a mere historical accident of grammatical habits of separating a subject and a predicate in our sentences, which later created an illusion that there really is a *self* to which we can refer.

In contemporary analytic philosophy, Daniel Dennett represents the skeptical, eliminativist view of a single self as the subject in consciousness. Based on current neuroscientific discoveries, Dennett argues against what he calls "the Cartesian Theater model." The Cartesian Theater model projects a single self as the observer of one's flow of consciousness, the *I* who is both the Cartesian thinker and the one who engages in self-inquiry. However, Dennett argues, neuroscience has discovered that "there is no single point in the brain where all information funnels in," and "there is no observer inside the brain" (Dennett 1992, 103). The correct picture of our consciousness is to think of parallel information processing tracks in the brain, producing constantly revised "drafts" that interpret and reinterpret what we are experiencing. Hence, we do not have a single narrative of our conscious life that belongs to a single agent; what we have instead are multiple drafts undergoing continuous "editorial revision."

While there are some contemporary analytic philosophers, like Dennett, who are skeptical of the self, there are far more who affirm its existence and seek to clarify its nature. Sydney Shoemaker (1986) agrees with Hume that there is no such thing as an introspective sense impression of the self. He argues that introspection involves relational knowledge that stands between an act and an object, but the self, being a mental *subject*, cannot itself be the object of introspective awareness. In other words, the self as the *I* cannot at the same time be the *me* of the same self. However, the word 'I' is more fundamentally used as subject than as object, according to Shoemaker. 'I' refers to the subject of statements and each person's system of reference has the person himself as an "anchoring point" (Shoemaker 1968, 567). Roderick Chisholm (1969) argues that Hume's mistake begins with his using "perception" as a mode of self-awareness, since the self is not supposed to be a perceivable object, but *that which* sees, hears, loves, or hates. He also points out that Hume's argument is self-defeating because in Hume's self-report, there was already an *I* who "stumble[s] on some particular perception or other, of heat or cold, light or shade, love or hatred, pain or pleasure" (Hume [1739] 2000, 252). This Hume-person who made the discovery is the

subject *I* who apprehends all these perceptions in Hume's self-report. Therefore, that very Humean denial already proves an awareness of the self as subject. Even if the *I* cannot be an object of introspection, perception, or awareness, the *I as subject* seems indispensable.

Not only must there be an *I* who is the doer of deeds and the thinker of thoughts, but also there must be an unmistakable *I*, since an erroneous identification of the self is impossible. Wittgenstein argues that when 'I' is used as subject, there is no possibility of error in self-identification. If I say "I have a toothache," for example, it would be impossible that "I should have mistaken another person for myself" (Wittgenstein [1958] 1969, 67). Shoemaker (1968) also argues that when we use the word 'I' as the subject of our statements, we do not need to go through an identification process through which we identify ourselves as having the properties asserted in those statements. The reason is twofold: first, if we have to identify the self through some descriptive predicates, we must already possess a basic form of self-knowledge that we have these identifying features; second, identification goes with the possibility of misidentification, but in the case of the self, there is no possibility of misidentification. Therefore, basic self-knowledge is not based on identification of the self. For Shoemaker, self-knowledge or self-awareness comes in the form of self-predication: when one ascribes some particular predicates to oneself, such as "am hungry," "see a garden in front of me," "feel sad," "am in pain," one manifests self-knowledge or self-awareness. Shoemaker calls these special predicates *P**-predicates, "each of which can be known to be instantiated in such a way that knowing it to be instantiated in that way is equivalent to knowing it to be instantiated in oneself" (Shoemaker 1968, 565). Shoemaker argues that the self-ascription of these *P**-predicates is *immune to error through misidentification* – I cannot fail to identify *myself* when I use the word 'I' even though I could be mistaken about my beliefs about myself. If I ascribe to myself that I am in pain, for example, then I know that *I* am in pain. Self-knowledge is demonstrated in one's ability to use these *P**-predicates since using them presupposes having self-awareness. Our linguistic competence in self-ascribing *P**-predicates constitutes self-awareness.

In Shoemaker's analysis of self-knowledge, we see that self-reference is closely related to self-awareness. One could argue that our sense of the self is manifested in the linguistic usage of the first-person pronoun 'I,' which permeates our thinking and speaking. Even a person who suffers total amnesia is able to report: "*I* don't know who I am"; even a person with prosopagnosia would report, upon seeing her own image, "*I* don't recognize her." As long as a person can use the word 'I' in any statement, he or she has

a sense of the self however meager the information is. As Gareth Evans claims: "the essence of self-consciousness is self-reference" (Evans 1982b, 191). Such a linguistic habit of using the first-person pronoun is not a mere "historical accident," as Nietzsche has assumed, but an inevitable development of our language because of the way we think about and talk about ourselves. Hector-Neri Castañeda argues for such inseparability between self-reference and the self: "[A] correct use of 'I' cannot fail to refer to the entity to which it purports to refer . . . The first-person pronoun, without predicating a self-hood, purports to pick out a self *qua* self, and when it is correctly tendered it invariably succeeds" (Castañeda 1969, 161). Self-reference cannot fail to refer to the self, and this would then be the way to establish the self as subject. This is the classic Cartesian move, though not necessarily the Cartesian ego – an immaterial thinking substance. P. F. Strawson also claims that it seems to be "generally agreed" that an individual's use of 'I' is guaranteed against two kinds of failure: the failure of lack of reference and the failure of incorrect reference (Strawson 1994, 210).

G. E. M. Anscombe, however, disagrees. In "The first person," Anscombe argues that the word 'I' is not a referential term, and the use of 'I' does not guarantee that there is anything being referred to. She thinks that we derive a "grammatical illusion of a subject" from the seemingly self-referential nature of the first-person pronoun (Anscombe [1975] 1994, 159). Her argument can be defused by the view presented in John Perry's classic essay "The problem of the essential indexical." Perry analyzes 'I' as an essentially indexical term. An essential indexical depends on the context in which it is used to pick out the referent, and no other term could replace it without losing some of the explanatory force this term carries. The use of the first-person pronoun 'I' indicates a direct relationship between the speaker's conception and the speaker *herself*. His analysis of the referential nature of 'I' is externalistic and contextual: the usage of the 'I'-word itself, uttered by a particular speaker in a given context, secures the speaker as the referent. This view does not need to posit a Cartesian ego or any privately introspected self. Rather, it places the self in the midst of our language game and identifies the speaker as the subject. This is a self in the public sphere.

Galen Strawson (2009) points out a dual use of the word 'I' and its two associated conceptions of the self *as subject*: "when I think and talk about myself, my reference sometimes extends only to the self that I am, and sometimes it extends further out, to the human being that I am" (Strawson 2009, 31). The former use of 'I' refers to the subject of consciousness, as it is phenomenally presented to oneself, as how one conceives of oneself, how one is viewed or considered "from the inside"; the latter use of 'I' refers to

the whole human being, spatially and temporally located in the world, as the subject of physical or mental attributes, the agent of actions, the owner of moral and legal responsibilities. We can call the former "the phenomenal I," and the latter "the public I." Or we can also say that the former is the *subjective self* while the latter is an important part of the *objective self*. Thomas Nagel (1983) proposes a conception of the objective self, according to which the self (TN) is projected into the *centerless* world, singled out by a complete set of publicly identifiable properties and viewed from an impersonal standpoint. Nagel argues that each one of us has, or should have, an objective self at our core. Owen Flanagan (this volume) depicts a conception of the person as he or she really is, *from God's point of view*, as the person's "actual full identity." The self one conceives from the first-person point of view in all likelihood does not reflect the whole person truthfully or completely, seen from the impersonal standpoint or from God's point of view. There are definitely gaps between what we represent ourselves to be and what we really are. The self viewed from the subjective standpoint and the person from an objective standpoint may not match up, and here we see that the notions of *self* and *person* diverge. This divergence also leads to possible problems in self-knowledge: what we think about ourselves, in terms of our mental states or psychological attributes, could have missed the mark of what kind of person we truly are and what kind of beliefs we truly have.

The reliability of self-knowledge can be subject to the same Humean skepticism: how do we *introspect* our own beliefs and desires? The question is again whether our own mental states could be the *object* of our knowledge. Gareth Evans (1982b) suggests that the way to gain self-knowledge is not "looking within," but to look at the outside world to form a judgment about the world. When the subject wishes to know whether he believes that *P*, he "does not in any sense gaze at, or concentrate on, his internal state. His internal state cannot in any sense become an *object* to him. (He is *in* it.)" (Evans 1982b, 204, original italics). Evans advocates an externalist view of self-knowledge: "In fact, we only have to be aware of some state of the world in order to be in a position to make an assertion about ourselves" (Evans 1982b, 207–08). Alex Byrne (2005) follows Evans and proposes an externalistic epistemic method of self-knowledge. He argues that reliable self-knowledge is easily obtainable if one simply follows a "self-verifying" epistemic rule, which he calls BEL: If *p*, believe that you believe that *p*. This rule is self-verifying because if one follows it, then one's second-order belief about one's own belief will be true. In other words, one gains reliable self-knowledge by following this epistemic rule. The advantage of Byrne's

self-verifying principle is that it does not call for a special "internal mechanism" or an "inner sense" for our self-knowledge. As long as one perceives the conditions of the world and is rational enough to follow this epistemic rule, one *knows* what one believes.

Hume's concern about the introspectability of the self's presence and mental states can also lead to a question about the introspectability of the self's experiences, path of life, and *personhood* in general. Can we have a conception of the self that is not completely determined from the individual's point of view, the individual's self-ascription of beliefs, desires, and other psychological traits, and the individual's ownership of his or her deeds in memories? In other words, how do we establish an objective personhood that is not purely derived from the individual's phenomenal consciousness? The psychological account given by Locke seems insufficient as an account of personhood. Bernard Williams (1957) proposes that we use bodily continuity as a *necessary* condition of personal identity. He argues that the memory criterion cannot be divorced from the body criterion, since the only condition under which x has a veridical memory of y's doing A is that $x = y$, and for x and y to be the same person is to have the spatio-temporal continuity between their bodies. There may be imaginary cases of soul swapping or body switching, but if we go for a more realistic approach, according to Williams, "the facts of self-consciousness prove incapable of yielding the secret of personal identity, and we are forced back into the world of public criteria" (Williams [1957] 1999, 15). David Wiggins (1967) also uses spatio-temporal continuity as the criterion of personal identity. In recent literature, Eric Olson's (1999) account of personal identity appeals to the biological organism that human beings are, or biological continuity, as the criterion. He rejects the psychological criterion completely and suggests that a human person is just a living human animal. This view has been called "animalism." All these approaches can be seen as an attempt to establish personhood from an objective or at least *public* perspective, one that is not confined to the individual's consciousness.

The collected essays in this book continue these discussions in a new light. The first set of essays begins with the Humean denial. Some take up Humean skepticism about locating the self through our introspection while some aim to defeat it. The second set of essays deals with the issue of self-knowledge and the third set of essays explores the relation between personhood and one's consciousness.

David Rosenthal argues that Hume's denial of the self is based on an unfounded assumption that perception is the only means through which one could be aware of a self. Having an occurrent, assertoric thought about

one's self as being present is another way for one to be aware of it. Rosenthal's HOT theory defines 'conscious mental state' as the mental state in which one has a suitable thought that one is, oneself, in that state. If one thinks that one is in a mental state, say, thinking about which movie to go see, then one not only is conscious of the thought but also is conscious of the thought as belonging to oneself. The sense of self emerges as the owner of these conscious mental states. Uriah Kriegel argues that self-awareness is always minimally in the "peripheral" of our conscious phenomenal experience. A phenomenal experience is how the experience is an experience *for me*, the subject of the experience. Without the subjective aspect, the experience itself would not even be an experience for anyone. We may not be able to find the subject in our phenomenal experience, but the subject is always there in our consciousness. Jesse Prinz, on the other hand, elaborates on Humean skepticism and argues against the possibility of finding, in our phenomenology, the self as the subject. Prinz rejects what he calls "the phenomenal I," not what he calls "the phenomenal me." According to Prinz, we never can have an experience of ourselves as the subject of our experiences; what we have are just our experiences, our mental states, our perceptions of the world. We do not experience ourselves as the "owner" of conscious mental states as Rosenthal claims; nor do we have any qualitative experiences associated with a subject *for whom* those qualitative experiences are experiences as Kriegel claims. All three views can be seen as examining the self as subject; in particular, the *subjective self*, or the self viewed "from the inside."

On the *public self*, we have essays by John Perry and Lucy O'Brien. John Perry examines the nature of "self-beliefs" or "I-thoughts," and places their origin in the world. One gathers "information" about oneself by multiple means, some of which are publicly assessable. The self does not have to be a private entity, knowable only to the subject. Lucy O'Brien focuses on the self-consciousness involved in one's awareness of others' gaze, which she calls "ordinary self-consciousness." On O'Brien's conception of the self, the *I* is both the object of others' scrutiny and the subject who experiences the variety of emotions associated with being thus self-aware. There seems to be a double role for the self here: the first self is the publicly observable person (the public *I*) while the second self is the subject *I* who is imagining how others examine the pubic self.

The gap between the subjective self and the objective self leads to the issue of authoritative self-knowledge. The issue of self-knowledge is another theme explored in this book. If there is a unified self within our consciousness, then to know what one thinks, what one believes, or what one desires

should be the most reliable, immediate, and authoritative form of knowledge. The Cartesian claim of *infallible* self-knowledge has been taken to assume the subject's having a "privileged access" to the content of her own mind. However, even with the most honest intention, we do not have a complete grasp of ourselves. If personhood is construed purely subjectively, from the individual's self-conception, then it can deviate from the individual's real personality traits, moral character, true memories, and current intentions. Fred Dretske, Alex Byrne, and Eric Schwitzgebel examine different aspects of knowledge about one's self, including one's beliefs, desires, and moral attitudes. Dretske argues that "from the inside," the individual has no privileged evidence to the fact that she is thinking. Byrne suggests that we make judgments of *desirability* on the state of affairs in the world as a way to know our own desires. Schwitzgebel points out that our internal conception of ourselves frequently falls short of capturing our true selves. All three views can be seen as defending an externalist position on self-knowledge.

Finally, the last thread in this book is a reflection on the Lockean conception of personhood and persistence of the self in time on the basis of consciousness. Sydney Shoemaker defends a "neo-Lockean" theory of personal identity, using psychological continuity as the criterion for personal identity. Owen Flanagan appeals to William James' notion of *consciousness* and argues that personal identity should be based primarily on the Jamesian person. Even though the two views differ, they both appeal to the individual's consciousness to assign personal identity, and they both argue that the criterion cannot be determined purely "from the inside": for Shoemaker, it is the causal profiles of one's psychological properties; for Flanagan, it is the complete consciousness, the actual full identity not founded on the individual's autobiographical memories, that constitutes personhood.

SYNOPSIS OF CHAPTERS

In "Awareness and identification of self" David Rosenthal defends his higher-order thought (HOT) theory of self-awareness. Rosenthal addresses three concerns that the HOT theory may not be adequate as a theory of self-awareness. The first problem is that some of our thoughts about ourselves involve the so-called "essential indexical" – they are thoughts about one's referring to oneself with a first-person mode. These thoughts seem to be antecedent to one's having a higher-order thought about one's mental states; hence, some other account needs to be given to explain how we