

The Nature of Inner Experience

Kant on Self-Knowledge and Self-Formation

Katharina T. Kraus

The Nature of Inner Experience

Kant on Self-Knowledge and Self-Formation

Katharina T. Kraus

Contents

Preface

Introduction: Inner Experience, Psychological Persons, and Self-Formation

Chapter 1

Inner Sense as the Faculty for Inner Receptivity

Chapter 2

Temporal Consciousness and Inner Perception

Chapter 3

The Form of Reflexivity and the Expression of Self-Reference

Chapter 4

Consciousness of Oneself as Object: The Logical “I” and the Psychological “I”

Chapter 5

The Idea of the Soul as the Guiding Thread of Inner Experience

Chapter 6

The Demands of Reason: Self-Formation and Self-Knowledge

Bibliography

Index

Contents (Details)

Preface

Introduction: Inner Experience, Psychological Persons, and Self-Formation

1. From Inner Experience to the Formation of Psychological Persons
2. The Puzzle of Self-Reference: Parity or Disparity?
3. The Argument of the Book: Varieties of Objects and Varieties of Self-Consciousness
4. The Novel View of the Book: Self-Formation under the Idea of the Soul
5. Outline of the Chapters

Chapter 1 Inner Sense as the Faculty for Inner Receptivity

1. Kant's Broader Theory of Representation
 - The Basic Model of Representation: Reflexivity and Referentiality
 - Objective and Subjective Validity
 - Transcendental and Empirical Faculties
2. Inner Sense in the Historical Context
 - The Empiricist Tradition: Inner Sense as Perception of Mental Acts and Contents
 - The Rationalist Tradition: Empirical Consciousness Requires Self-Consciousness
3. Kant's Transcendental Account of Inner Sense in the *Critique of Pure Reason*
 - The Faculty of Sensibility: Inner and Outer Sense
 - Inner Sense and Empirical Consciousness in General
4. Inner Receptivity in *Anthropology* and *Critique of the Power of Judgment*

Chapter 2 Temporal Consciousness and Inner Perception

1. Perception and Synthesis
2. The Interactive Model of Perception
 - Outer Affection
 - Positing and Apprehension
 - Self-Affection
3. Transcendental Self-Affection and the Temporal Conditions of Perception
 - Two Functions of the Understanding
 - The Figurative Synthesis
 - The *A priori* Temporal Conditions of Perception
4. Empirical Self-Affection and Inner Perception

Chapter 3 The Form of Reflexivity and the Expression of Self-Reference

1. Transcendental Apperception and Transcendental Self-Consciousness
 - The Deduction Argument
 - The Unity Argument
 - The Notion of Transcendental Apperception
2. Transcendental Apperception as Act-Awareness
 - The Psychological Interpretation
 - Critical Assessment of the Psychological Interpretation
3. Transcendental Apperception as Capacity for Self-Ascription
 - Self-Ascription and Self-Reference
 - The Logical-Subject Account
 - Critical Assessment of the Logical-Subject Account
4. Transcendental Apperception as Form of Reflexive Consciousness
 - The General Form of Reflexivity
 - Transcendental Apperception as the Origin of Referentiality
 - The Unity Argument in the Transcendental Deduction, Reconsidered
 - Reflexive Consciousness and Self-Consciousness
5. The Expression “I think” and Self-Reference
 - The Apperceptive “I think” as a Mere Expression
 - “I think” (empirically taken) and Self-Reference

Chapter 4 Consciousness of Oneself as Object: The Logical “I” and the Psychological “I”

1. Object of Thought and Object of Experience
 - The logical “I” and the psychological “I”
2. The Logical “I” as an Object of Thought
 - The “I think” in the Paralogisms
 - From the Logical Point of View: Logical Predicates as the Rules of “I”-judgments
3. The Psychological “I” as an Object of Inner Experience
 - The Paradox of the Determinable Self
 - The Crux of Inner Experience: No Mental Persistence, but Personal Identity
 - The Psychological Point of View in the Third Paralogism

Chapter 5 The Idea of the Soul as the Guiding Thread of Inner Experience

1. Reason and Human Experience
 - Reason as the Faculty for Inferring
 - Pure Reason and Transcendental Ideas
 - The Regulative Use of Reason
2. The Idea of the Soul in the *Appendix to the Dialectic*
3. Transcendental Presupposition and Regulative Principle
 - The Difficulty of An Interpretation
 - Do Principles of Reason Describe A Given Reality?

4. Two Interpretations of the Idea of the Soul
 - Less Than a Real Thing: Against the Noumenal Interpretation
 - More Than a Heuristic Fiction: Against the Fictional Interpretation
5. The Idea of the Soul as the Regulative Principle of Inner Experience
 - The Transcendental Idea as an “Analogue of a Schema”
 - The Presupposition of a Systematic Mental Whole

Chapter 6 The Demands of Reason: Self-Formation and Self-Knowledge

1. The Self-Formation of Psychological Persons
 - Kinds of Self-Realization and Kinds of Self-Related Unities
 - The Idea of the Soul as the Form of Mental Life
 - Personal Identity and Personal Development
 - Cases of Mental Disunity
2. The Conceptualization of Psychological Phenomena
 - Three Principles of Concept-Formation
 - The Idea of the Soul as the Genus Concept of Inner Nature
 - The Idea of the Soul as the Placeholder for a Complete Self-Concept
 - A System of Psychological Predicates
3. A Normative Concept of A Person
 - The Normativity of Personhood as Inner Coherence
 - The Epistemic Model of Normativity: Epistemic Coherence
 - Self-Knowledge and the Possibility of Self-Deceit
 - Self-Knowledge and Inner (In-)Coherence
4. The Imperatives of Self-Formation and of Self-Knowledge

Bibliography

Index

Introduction

Inner Experience, Psychological Persons, and Self-Formation

1. From Inner Experience to the Formation of a Psychological Person

As the preeminent Enlightenment philosopher, Immanuel Kant is famous for emphasizing that each and every one of us is called to “make use of one’s own understanding without direction from another” (*Enlightenment*, 8:35). We are all called to make up our own minds, independently from the external constraints imposed on us, for instance, by social hierarchies, political orders, or religious institutions. In the face of this Enlightenment calling, much of Kant’s philosophy then reads as a manual for how to employ one’s mental capacities in the proper way. It provides detailed accounts of various mental faculties such as the senses, the understanding, reason, the imagination, the will, the faculty of desire, and the feelings of pleasure and pain, specifying their functions and the conditions of their proper employment. These mental faculties are supposed to be universally realized by each human being. Given his focus on a universal conception of the human mind, however, Kant tells us surprisingly little about what makes us unique individuals. What, for Kant, makes us the particular empirical persons we are and how do we come to know ourselves as such persons?

Even on closer inspection, it is not obvious which conception of a person is implicit in Kant’s Critical philosophy and how knowledge of oneself as a human individual is possible. On the one hand, a person must view herself as a mind endowed with mental capacities, such as the faculties for feeling, reasoning, and willing. On the other hand, a person must view herself as a being that is embedded in the spatiotemporal and causally structured empirical world, and that is characterized by various psychological features, including a unique psychological character and an individual course of mental life. Both these ways of viewing oneself are crucial for understanding oneself as a *psychological person*. This book offers an account of *psychological personhood* by unfolding, in accordance with the tenets of his Critical philosophy, his account of *empirical self-knowledge*, i.e., the knowledge that one has of oneself as a unique psychological person.

In Kant's account of personhood, then, a central role is played by the capacity to judge about one's own psychological features, that is the capacity for what Kant calls *inner experience*. Primarily, inner experience concerns a person's conscious mental states, such as occurrent sensory perceptions, thoughts, imaginations, feelings, desires and intentions for action, as well as memories of past states and their causal relations. Moreover, inner experience also concerns general psychological properties such as psychological dispositions, personality traits, social relationships, and moral, political, or religious values. Through inner experience, persons gain access to a range of psychological phenomena that make up their mental lives and which in turn determine who they are. Although inner experience has been neglected in the contemporary literature on Kant, I argue that, for Kant, it is a primary means by which we not only gain empirical self-knowledge, but also form ourselves as empirical persons. So in this book I defend two central theses.

First, for Kant, *inner experience is empirical cognition of oneself, not as a mere object, but as a psychological person.*¹ On my reading, Kant conceives of inner experience as analogous to the experience of mind-independent objects in some respects, yet as fundamentally different from it in other crucial respects. On the one hand, a person should be construed by analogy with external objects of experience insofar as the person's specific psychological features are embedded in the spatiotemporal, causally structured world and therefore give rise to a particular kind of experience, namely inner experience. On the other hand, a person fundamentally differs from *mere* objects of experience in that a person must also be construed as a mind endowed with distinct faculties for representation and with the ability for self-determination. Persons must be conceived of as determining – at least to some extent – who they are and what their psychological features are in the course of their mentally active life.

My second thesis is, then, that *psychological persons form themselves in the course of realizing their mental capacities under the guidance of a unifying idea, the idea of the soul*. So this book defends what I call the *self-formation view* of the psychological person. According to this interpretation, a psychological person is understood not as a self-contained entity that exists prior to the particular happenings

¹ Throughout this work, I employ the term "experience" in this Kantian sense, which I introduce in Section 2 below.

of one's mental life, but rather as an entity that first emerges through self-formation in the course of mental activity. More precisely, a psychological person organically evolves as a mental whole by unifying all mental activities and the states in consciousness that result from these activities, such as perceptions, beliefs, desires, memories, dreams, and feelings, under the guidance of an idea, viz. the idea of the soul. The central task of such an interpretation is to discern the conditions that make the formation of oneself as a psychological person possible.

I focus primarily on the experiential, rather than the agential, side of personhood. That is, I explore the nature of psychological persons insofar as they can know themselves through inner experience, rather than insofar as they act as rational agents in the world. Therefore, I examine the conditions of self-formation with regard to the conditions of inner experience, rather than with regard to the conditions of agency. For the former, I draw mainly on the resources of Kant's theoretical philosophy, whereas the latter would involve a close examination of Kant's practical philosophy. Despite confining myself primarily to the experiential side, I firmly believe that the conception of psychological personhood I am offering here is compatible with Kant's theory of agency and can be expanded in this direction in the future.

In sum, my interpretation has the following three characteristics:

- (1) It remains agnostic regarding the intrinsic nature of that which appears in inner experience, such as an underlying pre-existing soul or non-material substance.
- (2) It takes the idea of the soul – as the concept of a unified mental whole – to be a regulative guideline for understanding of one's inner experiences as representing psychological determinations of one and the same person – oneself.
- (3) It takes the idea of the soul to be a prescriptive principle of mental life that is practical efficacious with respect to the self-formation into a unified, integrated psychological person.

On this interpretation, an individual person is precisely the unique mental whole that is progressively realized through exercising mental capacities in accordance with the unifying principle prescribed by the idea of the soul. The starting point from which I develop this interpretation is the puzzle of self-reference, to which I now turn.

2. The Puzzle of Self-Reference: Parity or Disparity?

Empirical self-knowledge raises an intricate puzzle – a puzzle that indeed is a problem for any philosophical or scientific theory addressing it. On the one hand, self-knowledge is *reflexive* in that it points back to the subject who has the experience. On the other hand, self-knowledge *refers* to a particular individual, namely oneself, with specific psychological features. The puzzle thus concerns the issue of how subjects can represent themselves as objects, that is to say, how self-knowledge can be *self-referential* at all. It is not clear whether and, if so, how subjects are able to refer to themselves as objects, without thereby reifying or doubling themselves in illegitimate ways.

Let me expand on this thought. Self-knowledge seems to involve two ways of representing oneself: representing oneself *as subject* and representing oneself *as object*. In contemporary philosophy of mind, this issue has been reflected in the distinction between two kinds of self-consciousness, between “consciousness of oneself *as subject*” and “consciousness of oneself *as object*”.² Contemporary philosophers of language have specified two uses of the first-person pronoun “I”: the use of “I” as subject and the use of “I” as object.³ While this terminology is certainly helpful, it leaves the following two questions unanswered. Firstly, if I were to become the object of my own experience, what *kind of object* would I be for myself? Would I be a mind endowed with capacities and dispositions, a spatiotemporally embodied human being, the sum total of all my mental states, or rather something else? Secondly, if my self-knowledge is primarily concerned with the contents passing through my subjective consciousness, can these contents ever become items of knowledge that meets the standards of *objective validity*? That is, can I objectively know myself?

This study is driven by the belief that Kant’s Critical philosophy – as an enquiry into the necessary conditions of the possibility of experience – provides an exceptionally productive framework for examining these questions. In his *Critique of Pure Reason* (1781/1787; henceforth, *Critique*), Kant assigns a central role to the

² E.g., Cassam 1997, Carl 2014, Frank 2007, Longuenesse 2017.

³ E.g., Wittgenstein 1958, Shoemaker 1968. *Prima facie*, it is unclear whether and how these distinctions map onto Kant’s distinction between subject and object, though they certainly seem to address a similar set of phenomena. For an insightful comparison between these contemporary distinctions and Kant’s, see Kitcher 2000 and Longuenesse 2017.

thinking subject in the constitution of experience of objects and thereby conceives of experience as *empirical theoretical cognition of objects*, rather than as merely subjective sensation.⁴ A detailed analysis of Kant's account of mental faculties offers, I shall argue, crucial resources for resolving the puzzle regarding inner experience. Such an analysis will reveal the distinctive ways in which we relate to ourselves as objects, whilst at the same time acknowledging our nature as thinking subjects.

According to Kant's transcendental philosophy, *experience (Erfahrung)* is a kind of *empirical cognition (empirische Erkenntnis)*, that is, experience consists in sensation-based judgments about an object or an objective reality. Experience results from the mind's activity of bringing a multitude of sensations under empirical concepts and of combining those concepts into judgments. The two main faculties involved in experience are the faculty of sensibility and the understanding as the faculty to judge. *Sensibility (Sinnlichkeit)* gives us immediate access to objects of experience. It yields sensory *intuitions (Anschauungen)* of them in the form of time and space. The *understanding (Verstand)* applies general *concepts (Begriffe)* and thereby yields *judgments (Urteile)* about objects of experience. It does so in accordance with its basic forms, the so-called categories such as unity, substance, and causality. Both the forms of sensibility and the forms of the understanding are conditions of the possibility of experience that are universally shared by all humans.

The paradigmatic case of experience that Kant considers throughout the *Critique* is the cognition of material objects in space and time. The notion of inner experience is nonetheless a ubiquitous theme in the first *Critique*, as well as in other works from the same period.⁵ Kant construes inner experience – by and large – on analogy with outer experience. By “observing” ourselves through inner sense, as opposed to observing outer objects through outer sense, we are able to become aware of the momentary perceptions, emotions, and thoughts that we are empirically conscious of, as well as, arguably, of our long-term moods, passions, and standing attitudes. On the basis of this awareness, we are then able to make judgments about our mental states and their temporal relations, and arguably about more general

⁴ In this study, I confine myself to empirical theoretical cognition of objects, which I will just call “cognition”, if not stated otherwise.

⁵ E.g., CpR B277-279; A672/B700; Anthr 7:141-142, 7:161-162.

psychological properties, such as character traits. Our inner experience consists in such empirical judgments about our own psychological features.⁶

This conception of inner experience raises precisely the puzzle that I have stated above. In light of Kant's account of experience in the *Critique*, we can now ask: firstly, if inner experience concerns that which appears in my subjective empirical consciousness, can it then fulfil the conditions necessary for the cognition of an object? Secondly, can the subjective consciousness that I have of my own feelings, desires, and thoughts give rise to objectively valid cognition of myself? Kant himself is aware of this intricate puzzle, which is raised by the fact that "the I that I think is to differ from the I that intuits itself [...] and yet be identical with the latter as the same subject" (B155-156). But unfortunately Kant does not offer a clear account of his solution to the puzzle.

Despite its frequent recurrence in the *Critique* and other texts, the theme of inner experience has mostly aroused suspicion and perplexity among commentators regarding its nature and epistemic status, and only little interpretive consensus about it has been achieved. Most notably, there is a striking lack of discussion of it in the Anglophone literature, and also a surprisingly small amount of discussion in German or in other languages.⁷ When it is discussed, various commentators are skeptical about whether inner experience is able to play a significant role in Kant's Critical philosophy, and some even deny that inner experience can be theoretical cognition of an object at all.⁸

There are two main rival lines of interpretation of inner experience within Kant scholarship and beyond. One line of interpretation construes inner experience as cognition of a mental object on a par with the construal of outer experience as cognition of physical objects. Another denies that such an objectual or objectified grasp of oneself is in any way possible and claims that self-knowledge, if one can speak of it at all, must be knowledge of an entirely different kind. In the history of

⁶ He also uses other phrases, each emphasizing a different aspect, such as "determining my existence in time" (see CpR Bxl, A35/B53, B157n, B430-431), "cognizing myself as I appear to myself" (see B68, B155, B157-158, A337/B394), and "connecting my inner appearances and actions" (see B156, A672/B700, A683/B711).

⁷ The notion of inner experience is rarely discussed or altogether absent from studies of Kant's accounts of the mind, the self, and empirical psychology, such as Ameriks 2000, Sturm 2009, Kitcher 1990, 2011, Wuerth 2014. Exceptions include Mohr 1991, Emundts 2007, Frierson 2014, and Dyck 2014.

⁸ For references, see my discussion of the disparity view below.

reception of Kant's thought, both these approaches have had persuasive defenders and both are still alive today. The tendency towards the objectification of the mental has been pursued predominantly by Neo-Kantian and naturalist interpretations of Kant's philosophy; the subjectivist line of interpretation has been defended mainly by German Idealists, phenomenologists, and existentialists.⁹ This leads me to distinguish, in a more systematic way, two interpretive approaches with respect to Kant's conception of inner experience, both of which are problematic: the *parity view* and the *disparity view*.

According to the *parity view*, inner experience is construed in the same way as outer experience and so, like it, as empirical cognition of an object. This view appeals to an alleged structural parallel between inner and outer sense and holds that inner and outer experience are determined through the same set of formal conditions, despite some specific differences, such as the non-spatiality of mental states and the lack of the attractive and repulsive forces of matter in the case of inner experience.¹⁰

In contrast, another diverse set of interpretations emphasizes the disparity between inner and outer experience and therefore rejects the identification of inner experience with the empirical cognition of an object in the proper sense. Although these interpretations come in a wide variety of flavours, I subsume them under the common label *disparity view*. Some defenders of the disparity view argue that inner experience is more fundamentally dependent on the consciousness of oneself qua thinking subject than is outer experience. Therefore, they conclude that inner experience amounts not to cognition of a mental object but to some sort of empirical awareness of one's own mental activity – an empirical self-awareness that is conditioned only by the features of reflexive thought.¹¹ Others detect an apparently insurmountable disparity between inner and outer experience on the grounds that inner experience allegedly lacks a referent, since no object is given to inner sense that could be determined by the category of substance.¹² For a third group, Kant's notion

⁹ For the former, see Wundt 1888, 1902, Cohen 1885, Natorp 1912; for the latter, see Fichte 1796/99, Husserl 1900, 1931, Heidegger 1929.

¹⁰ A strong tendency towards the parity view can be found in Vogel 1993, Frierson 2014, Chignell 2017; for closer discussion, see Kraus 2019.

¹¹ Such views have been stated by Henrich 1994 and Keller 1998; see also Kitcher 2011, Brook 1994, Ameriks 2000, Rosefeldt 2000, Hatfield 2006, Schmidt 2008, Longuenesse 2017.

¹² Strawson 1966, Mischel 1967, Gouaux 1972, Washburn 1976, Förster 1987, Nayak/Sotnak 1995, Westphal 2004, Friedman 2013, Schmitz 2013.

of inner experience can be meaningfully discussed only as an aspect of practical deliberations in the context of agency. For them, the only possible role for inner experience is *not* to yield theoretical cognition about oneself, but only to contribute to reflection on practical matters.¹³ Furthermore, Kant's apparent denial of proper scientific status to empirical psychology in the *Metaphysical Foundations of Natural Science* (1785; henceforth *Foundations*) has led many commentators to think that Kant must reject the very possibility of any theoretical knowledge of psychological phenomena. Kant's reservations about the mathematisability and objective observability of inner states have been taken as evidence against the possibility of inner experience as empirical cognition.¹⁴

The novel account that I develop in this study proposes a more nuanced understanding of inner experience as empirical cognition, which is able to answer the worries of the disparity theorists. In line with the parity view, I shall argue that inner experience is empirical cognition of a person's psychological features. Yet the special nature of persons, which gives rise to some of the disparities pointed out above, will warrant a crucial qualification: inner experience is only *analogous* to experience of spatiomaterial objects. The most important difference here is that inner experience is not cognition of a persistent substance given to inner sense; rather, the object of inner experience is first formed in the course of one's mental activity under the guidance of the unifying idea of the soul. Hence, the account of inner experience offered here – in combination with the self-formation view of psychological personhood – preserves conceptual coherence with Kant's notion of outer experience and yet remains sensitive to the distinctive systematic concerns that arise from the puzzle of self-reference.

¹³ This view is expressed by Sturm 2001, 2009:205-260 and Emundts 2017. Assertion of the primacy of the practical over the theoretical role of inner experience can also be found in Makkreel 2001 and Cohen 2009.

¹⁴ See MFNS 4:471. E.g., Gouaux 1972, Mischel 1967, Washburn 1976, Leary 1978, Schönrich 1991, Friedman 2013, Hudson 1994, Klemme 1996, Pollok 2001, Makkreel 2003. For a discussion of the scientific status of empirical psychology, see Kraus 2018.

3. The Argument of the Book: Varieties of Objects and Varieties of Self-Consciousness

What is at stake in the dispute between parity and disparity views is whether we can discern an “object” of which inner experience could achieve empirical cognition. The central question is then: what, if anything, do we cognize in inner experience? Reading Kant’s texts, the following candidates may recommend themselves: the thinking subject (*denkendes Subjekt*), the self (*Selbst*), the mind (*Gemüth*), the soul (*Seele*), a mere collection of inner appearances (*innere Erscheinungen*), or the whole embodied human being (*Mensch*). A central element of inner experience is the term “I” (and its cognates), as Kant himself states in a note:

All my inner experience is a judgment in which the predicate is empirical and the subject is *I*. (Refl 5453 18:186, my translation)

Accordingly, inner experience typically consists in what one may call *I-judgments*, i.e., judgments that can typically be expressed by statements of the basic form “I φ ” (and derivations thereof), whereby the subject term is always the first-person pronoun “I” and “ φ ” a mental predicate. Examples include the self-ascription of occurrent mental states, such as “I see a red ball”, “I think that $2^2 = 4$ ”, or “I am in pain”; as well as the self-ascription of more general, temporally extended properties and characteristics such as “I am a kind and generous person”, “I value honesty and authenticity”, or “I aim to become a successful piano player”.¹⁵ We can now ask more precisely: what, if anything, does the term “I” refer to in such I-judgments, and can it be used to represent an object of experience?

What exactly, on Kant’s view, “I” denotes, if employed in such I-judgments, can be understood only within his more comprehensive theory of representation, which spells out the basic relation of a subject towards an object in terms of a variety of mental faculties and their characteristic forms. Based on this account of mental faculties, Kant distinguishes several kinds of objects and, as I shall argue, several corresponding kinds of self-consciousness some of which involve the use of the term

¹⁵ In this study, I use the notions of *psychological features* and *psychological determinations* to refer both to relatively short-lived, frequently fluctuating *mental states*, which include in particular occurrent representations passing in empirical consciousness, and relatively stable, long-term *psychological properties* that characterize the whole person, such as character traits. In some cases, a clear distinction between states and properties may not be possible or useful. One aim of this study is to develop a unified account of the psychological person, accommodating the possibility of both mental states and personal properties.

“I”. Given Kant’s account of empirical cognition as involving both sensibility and the understanding, we can distinguish at least three fundamental ways in which a human subject can relate to an object, corresponding to the following three types of representation:¹⁶

- (R-i) via sensible representations, viz. *intuitions*, through which a particular object is given;
- (R-ii) via intellectual representations, viz. *concepts*, through which an object is thought;
- (R-iii) via *empirical cognition*, through which an object is cognized on the basis of (R-i) and (R-ii).

So, for Kant, empirical cognition necessarily consists of two components: empirical intuitions, through which an object is given, and empirical concepts, through which an object is thought.¹⁷ The necessary conditions of possibility of cognition – or what Kant calls its *transcendental conditions* – therefore include time and space as the forms of sensibility and the categories as the forms of the understanding. If a representation satisfies these transcendental conditions, then it is objectively valid, i.e., it can be assessed as to its truth with respect to the object and independently of the particular mind that has the representation.

Corresponding to these three types of representation, we can now distinguish three kinds of object:

- (O-i) the *object of the senses* (commonly called “appearance”, *Erscheinung*);¹⁸
- (O-ii) the *object of thought* (sometimes called “thought-entity”, *ens rationis*, *Gedankending*);¹⁹
- (O-iii) the *object of experience*, i.e., an *object of empirical cognition* (commonly called “object”, *Gegenstand*).

¹⁶ I here use the notion of object in the most general sense as something to which we relate by means of conscious representations. Even a mere representation can count as an object in a minimal sense (see A189-190/B234-235), as well as an action that is represented by a desire (CpracR 5:21).

¹⁷ See A51/B75-76 and A92/B125.

¹⁸ E.g., “The undetermined object of an empirical intuition is called appearance” (A20/B34), see also Bxxxvi.

¹⁹ A thought-entity, strictly speaking, is an object of *mere* thought only, i.e., “the object of a concept to which no intuition that can be given corresponds” (A290/B347).

Kant provides complex arguments for how these fundamental ways of relating to objects can be integrated into a single account and in particular for how intuitions are amenable to the conceptual conditions of thought and judgment. Some of these arguments will be relevant to the course of my later analysis. Important at this stage is the insight that taking something as an object does not necessarily mean representing it as an object of experience. We are now in a position to refine the central question and ask whether inner experience satisfies the transcendental conditions for experience and thus for theoretical cognition of an object. If this were not the case, then the inner domain for Kant would only amount to a merely subjective awareness of mental states in a more or less indeterminate manner, and not to experience in his distinctively demanding sense. Unfortunately, Kant remains ambivalent with regard to the existence and status of the object of inner experience. He appears to subscribe to two seemingly conflicting claims:

(K1) *No object is given to inner sense that corresponds to an empirical inner intuition and that subsequently can be cognized in inner experience.*²⁰

(K2) In inner experience, I cognize myself *as an object*.²¹

The former claim appears to support the disparity view, whereas the latter claim seems to support the parity view.

In order to resolve this conflict, the book sets out to examine which types of representations and which kinds of objects are involved in inner experience. The main argument of the book thus consists in discerning the conditions of different types of self-consciousness, i.e., of types of representations a subject can have of itself. Corresponding to the notions of object defined above, I distinguish the following types of self-consciousness:

(SC-i) consciousness of one's inner appearances, received by inner sense;

²⁰ E.g., "Inner sense, by means of which the mind intuits itself, or its inner state, gives, to be sure, no intuition of the soul itself, as an object" (A22/B37); moreover, A107, B275, A350, A381, B412.

²¹ E.g., "I as intelligence and *thinking* subject cognize myself as an object that is *thought*, insofar as I am also given to myself in intuition, only, like other phenomena, not as I am for the understanding, but rather as I appear to myself" (B155). "Cognition of myself" is mentioned at Bxl, B139, B158, B277-9, B400, B431; Anth 7:142, 161, Prol 4:336. The "object of inner sense" is specifically mentioned at A342/B400 ("I, as thinking, am an object of inner sense, and am called 'soul'") and A368 ("the object of inner sense (I myself with all my representations)"), and A357, A385; B403; B415; A443/B471; A846/B874; CpracR V95; MFNS 4:467; 4:542; Refl 6313 (18:614); Anthr 7:142.

(SC-ii) consciousness of oneself as an object of thought, referred to by the term “I”;

(SC-iii) consciousness of oneself through inner experience, expressed in empirical I-judgments.

All these types of self-consciousness – as I will show – have to be classified as *consciousness of oneself as object*. They have to be distinguished from another, still more fundamental type, namely

(SC-0) consciousness of oneself as thinking subject, via transcendental apperception.

I will call this also *transcendental self-consciousness*, since it defines – as I will argue – the general form of reflexivity that any conscious representation must display, regardless of its specific type. Transcendental self-consciousness is thus a condition of possibility of *all* empirical consciousness.

Let me add a further note of clarification with respect to Kant’s doctrine of transcendental idealism. The notion of object is ambiguous in yet another way in that it can “be taken in a twofold meaning, namely as appearance or as thing in itself” (Bxxvii). Only appearances can be determinable objects for a subject, whereas things-in-themselves transcend the bounds of our senses and cannot be determined through experience. My aim here is neither to defend an interpretation of Kant’s transcendental idealism, nor to give an account of inner experience in terms of the thinking thing-in-itself (or noumenon) that may be believed to appear in inner experience. Rather, I focus on those notions of object that are relevant for an analysis of inner experience, according to which an object in the most general sense is understood as something that can be represented and to some extent determined by a subject.²²

²² Etymologically, the English word “object” (as well as the German *Objekt*) is derived from the Latin *objectus*, literally meaning “that which is thrown or put before or up against” someone, namely the subject; a similar notion is implied by the German *Gegenstand*, literally meaning “that which stands up against”. In this sense, I understand the notion of an object primarily in relation to some subject, i.e., it is an *object for a subject* (see Gardner 1999:39).

4. The Novel View of the Book: Self-Formation under the Idea of the Soul

As a solution to the puzzle of self-reference, this book argues that in inner experience we cognize ourselves *not* as *mere* objects of experience, since we are not given to ourselves as objects in the first place. Rather, by acknowledging our nature as representing subjects, we must understand ourselves as psychological persons. Thus, it will be argued that in inner experience we cognize ourselves only qua cognizing our psychological determinations. In order to cognize ourselves qua psychological determinations, we nonetheless need to be guided by a unifying concept by which we come to understand ourselves as being identical throughout our mental lives. For outer objects, the category of substance is precisely the unifying concept that makes possible the experience of something persistent. For psychological persons, I argue that the *idea of the soul* substitutes for the category of substance, as well as for other categories that cannot be properly applied in the case of inner experience. This idea makes possible the unity and identity of persons through time. Hence, this book develops a novel account of the very “object” of inner experience, which accommodates parities as well as disparities between inner and outer experience, by offering a novel account of psychological personhood.

This book defends a novel account of personhood: a *self-formation view* based on the idea of the soul. The idea of the soul is an idea of reason, to which no corresponding intuition can be given. That is, we cannot intuit ourselves as souls, though we may very well conceive of ourselves as such. Reason, for Kant, is the highest intellectual faculty, whose primary function is the capacity for inferring, i.e., for drawing conclusions that go beyond that which is immediately given in experience. Assigning a central role to the idea of the soul within Kant’s Critical theory of experience may sound perplexing to many readers. The Paralogisms in the *Critique* are commonly read as a devastating criticism of the metaphysical theories of his day, including rationalist theories of the human soul. In consequence, Kant is often taken to deny us any theoretical knowledge of the human soul, including properties such as substantiality, simplicity, and non-materiality. Yet there is ample evidence that Kant never abandons the notion of the soul altogether, nor does he outright reject any meaningful employment of it. Even in his most mature published works, such as the three *Critiques*, and in his late lectures on *Metaphysics* and *Anthropology*, Kant still makes some positive use of this notion in the context of

theoretical reason and psychology – a fact that is rarely discussed in contemporary Kant scholarship or appreciated as an interesting and illuminating aspect of his Critical philosophy.²³ The self-formation view of psychological personhood recovers the idea of the soul by showing it to be not only compatible with Kant’s Critical philosophy, but also extraordinarily philosophically productive for accounts of personhood and inner experience. The crucial idea is that the idea of the soul – as the concept of a mental whole – defines the unifying form of mental life.

This interpretation is inspired by a claim that Kant puts forward in the *Appendix to the Dialectic*. There, after having diagnosed the misuse of reason common to many rationalist philosophies, Kant proposes a positive use of reason in terms of regulative principles of experience. Let me explain how the idea of the soul serves as the “guiding thread of inner experience” (*Leitfaden der inneren Erfahrung*), according to the following passage:

Following the idea [of the soul] as principle[], we will first (in psychology) connect all appearances, actions, and receptivity of our mind (*Gemüth*) to the guiding thread of inner experience *as if* the mind (*Gemüth*) were a simple substance that (at least in life) persists in existence with personal identity, while its states – to which the states of the body belong only as external conditions – are continuously changing. (CpR A672/B700).²⁴

The notion of the “guiding thread of inner experience” indicates a certain regulative – rather than constitutive or determining – status. This can then be spelled out by appeal to Kant’s well-known “*as if*”, which indicates a hypothetical, rather than an assertive use of the idea. So here Kant endorses the following conditional: if one’s mind really were a persistent simple substance, enduring across change, then the various particular acts and representations grounded in it would be connected on a robust metaphysical basis. For, if the antecedent were true, all of one’s mental acts or representations would simply be states that *one* simple substance (the mind) would be in at various times. Now Kant does not think we can assert the antecedent of this conditional as true without qualification. Yet he acknowledges a weaker sense in which we must assume the antecedent as a hypothesis, and indeed, in many contexts,

²³ Exceptions include Klemme (1996), Hessbrüggen-Walter (2004), Serck-Hanssen (2011), Dyck (2014), Wuerth (2014), and Falduto (2014:24-33).

²⁴ This central role of the idea of the soul for inner experience is rarely acknowledged in the literature. Two recent exceptions are Wuerth (2014), who tries to restore an ontological interpretation of the soul, and Dyck (2014:199-225), who correctly points out the role of the idea for investigations of inner appearances, but denies that these investigations amount to *cognition* of mental states.

proceed as if it *were* true: that is, as if our minds were persistent simple substances that ground all our mental actions and representations. Employed in this weak sense, the idea of the soul serves two roles.

With regard to inner experience, the idea of the soul defines a *transcendental presupposition* that is indispensably necessary to first understand oneself as a psychological person and which nonetheless can be assumed only as a *regulative principle*. Only by presupposing the idea of a mental whole that underlies all inner appearances can I cognize my inner appearances as the *mental states* of one and the same person – myself – without thereby cognizing or determining myself as a persistent mental substance.

In turn, the idea of the soul serves as the prescriptive principle of the self-formation of an individual person. The idea of the soul provides the unifying principle by which a subject combines all mental activities and all appearances in consciousness into a mental whole. That is, I make myself actual as a unique psychological person whose mental life unfolds in time by realising my mental capacities in accordance with the idea of the soul. The idea of the soul prescribes the form that gives unity to my mental life as a whole. This self-formation view does not assume that one's personhood is grounded in the objective reality of a soul or thinking substance that exists independently of and prior to the representational activities of the mind. Nor does it assume that one's psychological reality is the result of a merely subjective, perhaps arbitrary construction out of one's conscious representations. Rather, a psychological person is precisely the unique mental whole that is progressively realized through exercising mental capacities in accordance with the unifying principle prescribed by the idea of the soul.

In sum, the idea of the soul fulfils a double purpose in Kant's Critical philosophy. Firstly, it serves as a prescriptive principle of reason for the purpose of self-formation as an individual person. It serves, secondly, as a transcendental and regulative presupposition for the purpose of self-knowledge – that is, for the comprehension of oneself as *this* unique person through the course of inner experience.²⁵ The novel interpretation of Kant's account of psychological personhood

²⁵ My conception of self-formation shares similarities with *self-constitution* accounts, such as that of Korsgaard (1999, 2008), who pursues a *self-constitution* account with respect to Kant's view of practical reason and moral agency. While there are certainly similarities between her view and mine, the two accounts are for the most part complementary to one another. Korsgaard develops her view drawing broadly on both the Kantian and the Platonic-Aristotelian tradition. The two interpretations

that this book offers will, I believe, eventually foster a better understanding of Kant's account of practical personhood, as well.

5. Outline of the Chapters

This book consists of six chapters, each investigating a different aspect of inner experience in accordance with the different ways in which human subjects can relate to themselves: through sensible intuitions of which we become conscious in perception, through concepts combined in thoughts, and finally through inner experience, construed as a kind of empirical cognition and guided by an idea of reason. To this effect, Chapters 1 and 2 are devoted to developing accounts of inner sense and inner perception; Chapter 3 examines transcendental self-consciousness and the general form of reflexivity; and Chapter 4 carves out a general conception of consciousness of oneself as object and introduces the distinction between two uses of the term "I": the logical "I" and the psychological "I". Chapters 5 and 6 are devoted to exploring the idea of the soul as the "guiding thread" for inner experience and finally as the fundamental principle of self-formation for psychological persons. In what follows, I offer a brief outline of each chapter.

Chapter 1, "Inner Sense as the Faculty for Inner Receptivity", sets the stage for my interpretation by introducing Kant's broader theory of representation and by defining two pairs of concepts that will guide my analysis: reflexivity and referentiality, on the one hand, and objective and subjective validity on the other. Through an examination of the historical context, the chapter then develops an account of inner sense as a transcendental faculty of sensibility according to the Transcendental Aesthetic of the first *Critique*. In doing so, it gives preliminary accounts of central concepts, including *affection*, *sensation*, *appearance*, *intuition*, *perception*, and *experience*. As a result, the chapter suggests that – on analogy with outer sense – inner sense receives inner appearances and yields a specific type of representation, i.e., inner intuition, according to its specific form, i.e., time. Yet the full argument for this claim will be put forward only in Chapter 2. Chapter 1 finally suggests a broader notion of inner receptivity as the susceptibility to all mind-internal

focus on different problems in that hers focuses on the moral agent, rather than the psychological person or the idea of the soul, though Korsgaard considers the arguments Kant provides in the Postulate of the Immortality of the Soul (CpraR 5:122-123). Moreover, she does not trace the notion of self-constitution back to a *regulative* use of the idea of the soul.

causes, whereby it incorporates insights regarding the faculties for desire and feeling from Kant's *Critique of the Power of Judgment* and his *Anthropology*.

Chapter 2, "Temporal Consciousness and Inner Perception", offers an interpretation of inner perception as the perception of distinctively inner appearances by drawing on resources from the Transcendental Aesthetic and the Transcendental Analytic (mainly the A-Deduction) of the first *Critique*. The chapter develops an interactional model of perception with three constitutive aspects: (i) affection through outer sense, (ii) synthesis of apprehension through the active faculties of the mind, viz. imagination and understanding, and (iii) self-affection through inner sense. Each of these constitutive aspects is shown to define a formal and a material condition of perception in general. By carving out the notion of transcendental self-affection, i.e., the *a priori* determination of the form of inner sense through the understanding, the chapter derives the temporal conditions of perception in general. Finally, the general model is applied to the case of distinctively inner perception. It is argued that inner perception is empirical consciousness of mental states, based on the empirical self-affection of inner sense.

Chapter 3, "The Form of Reflexivity and the Expression of Self-Reference", explores the role of transcendental apperception for inner experience according to the Transcendental Deduction (B) of the first *Critique*. The chapter argues that transcendental apperception gives rise to transcendental self-consciousness, i.e., consciousness of oneself as thinking subject, by showing the insufficiencies of two alternative accounts of transcendental apperception defended in the literature, namely the psychological view and the logical view. The chapter then goes on to argue that transcendental apperception defines the *general form of empirical consciousness*, i.e., *reflexivity*, which is the most general condition that any conscious representation must fulfil. The chapter concludes by arguing that the phrase "I think", which must be able to be attached to all meaningful representations, serves as an *expression* of this general form, if viewed transcendentially, and as an *expression of self-reference* of an individual thinker, if viewed empirically.

Chapter 4, "Consciousness of Oneself as Object: The Logical 'I' and the Psychological 'I'", examines two ways in which one can conceptually represent oneself in judgments, in light of the results of the Paralogisms (in the Transcendental Dialectic of the first *Critique*). Through the logical "I", one represents oneself as an object of thought. The logical "I" defines the way in which any thinking subject must

refer to itself in thought. Hence, the logical “I” determines a set of logical predicates that are shown to be conditions of I-thoughts in general. By contrast, the psychological “I” is used to refer to oneself as a psychological person in empirical judgments and under the conditions of sensibility, i.e., under the temporal conditions of perception (which were derived in Chapter 2). Yet it is not immediately clear how these temporal conditions can apply to this case. Through closely reading the Paralogism of Personal Identity, the chapter concludes that we cannot cognize ourselves as persistent empirical substances throughout time, nor can the merely logical “I” capture the temporal dimension of personhood. Rather, inner experience requires a different kind of sensible explication of the category of substance.

Chapter 5, “The Idea of the Soul as the Guiding Thread of Inner Experience”, explores the idea of the soul provided through reason as the regulative principle of inner experience (as discussed in the Appendix to the Transcendental Dialectic of the first *Critique*). The chapter argues that the idea of the soul regulatively guides the unification of inner appearances into a mental whole such that those appearances can then be cognized as the mental states of such a whole, without cognizing the whole as such. More specifically, I show that the idea of the soul serves as an “analogue of a schema” (A655/B693) for all those schemata that cannot be applied to inner appearances, including the schema of substance as persistence. I develop my view in contrast to two rival interpretations: the noumenal view, which conceives of the soul as a noumenal substance, and the fictional view, according to which the soul is a mere fiction. So this chapter finally establishes the first central thesis that the idea of the soul makes possible inner experience as empirical cognition of oneself construed as a systematic mental whole.

Chapter 6, “The Demands of Reason: Self-Formation and Self-Knowledge”, completes Kant’s account of psychological personhood by showing that the idea of the soul defines the form of mental life through which we first become individual persons. So this chapter finally establishes the central thesis of my self-formation view, namely that psychological persons first form themselves as such in the course of realizing their mental capacities under the guidance of a unifying idea, the idea of the soul. It shows that the principles of systematicity defined by the idea of the soul express a normative demand with regard to both the capacity for self-formation and the capacity for self-knowledge, i.e., knowledge of oneself as unique psychological person. Moreover, the chapter argues that this demand is normative for the basic

conceptualizing and cognizing activities of inner experience. We are bound to conceptualize all inner appearances and all more general psychological properties in accordance with a system of psychological predicates, the highest genus concept of which is the idea of the soul. Furthermore, the idea of the soul is shown to serve as the placeholder for each complete self-concept that refers to an individual person and contains all psychological predicates that apply to this person. Finally, the chapter explicates the intrinsic normativity of personhood in terms of a demand for inner coherence among all mental faculties. This demand will then be articulated in the form of two imperatives, which are inextricably linked: the *imperative of self-formation* and the *imperative of self-knowledge*.