SELF-REFERENCE AND SELF-AWARENESS

Edited by

Andrew Brook

and

Richard DeVidi

Carleton University, Ottawa

Table of Contents

Acknowledgments

Introduction

Some intricate, very subtle differences between awareness of self and awareness of other things and the relationship of these differences to some equally intricate features of reference to self using first person pronouns ('I', 'my', 'me' and 'mine') and their third person equivalents (for example, 'he himself') have been the subject of a clearly delineated and largely self-contained dialogue within analytical philosophy for about thirty-five years now. The investigation began with seminal papers by Hector-Neri Castañeda (1966, this volume; 1967; 1968),took on important new dimensions in the work of Sidney Shoemaker (1968, this volume;1970) and led to further classical contributions by Perry (1977; 1979, this volume) and Evans (1981; 1982, this volume). It continues to inspire important new work to this day. Shoemaker called his paper 'Self-reference and self-awareness'. It would be hard to think of a more perfect title for the topics under discussion in this volume, so as well as republishing his seminal paper, we have also gratefully borrowed his title.

Three ideas have been central to this dialogue. They are that:

1. Certain uses of 'I' and cognates are ineliminable (the idea of the essential indexical)[1]

2. In certain situations at least, we are immune to error through misidentification of another as oneself (the idea of immunity to error through misidentification with respect to the first person)

3. When we refer to ourselves using 'I' and cognates, often or always it takes place without us identifying anything as oneself via properties that one has ascribed to the thing. (the idea of self-reference without identification)

These three ideas in various configurations have been central to the whole dialogue, from its creation (or recreation) in the 1960s to the present.

---

[1]Note that we do not refer to 'I', etc., as *indicators.* This may seem a bit odd but whether these pronouns and their cognates *are* indicators is a matter of debate. For example, one of Castañeda' three seminal papers is called 'Indicators and Quasi-Indicators'. He *distinguishes* the third person equivalents of first person pronouns such as 'I' from indicators. Because they depend on other indicators and/or referring expressions, he says that they are really quasi-indicators. It is an indication of the intricateness and complexity of the recent work on self-reference that even something as basic as this could be a matter for debate.

Creation or recreation? At the beginning of his 1966 paper, Castañeda says that his topic is "almost brand new". Though his paper is without question highly original and of first rate importance, Castañeda's claim is an exaggeration. No idea in philosophy (and precious few anywhere else) is ever constructed entirely from whole cloth. With respect to both self-awareness and self-reference in particular and indicators in general, that Frege was an important precursor is well-known, especially his (1918/19). The Wittgenstein of *The Blue and Brown Books* was another, as has been clear at least since Shoemaker's paper. It is less well-known that Kant also discussed some of the peculiarities of self-reference and self-awareness, indeed anticipated claims about them that had to wait 'til the work we've just introduced to get a full articulation.

In this collection, we start by exploring two of the great precursors of the recent dialogue on self-reference and self-awareness just mentioned, Kant and Frege. (There is no separate paper on Wittgenstein, the third, for reasons we will give in a moment.) Next, four of the classic contributions just mentioned are reprinted. The final section offers five examples of more recent work written for this volume. We will try to say something by way of introducing these papers but there are limitations on how much one can do. Because they are dense and intricately argued, there is no substitute for reading the papers themselves.

The papers on Kant and Frege that make up Section I are by the two editors. Andrew Brook argues that Kant not only anticipated but also clearly articulated the phenomenon that Shoemaker calls self-reference without identification, that is to say, that one can refer to oneself, and to refer to oneself *as* oneself, without (otherwise) identifying oneself. There are also indications that he was aware of something very much like the contemporary notion of the essential indexical. Brook argues, moreover, that, unlike contemporary theorists at least up to Peacocke (this volume), Kant had at least the makings of a theory to explain these peculiar features of self-awareness. His theory also yields insight into which references to oneself are and which are not immune to the error of identifying someone else as oneself (though Kant seems not to have thought of anything like this notion of immunity to error itself). Peacocke also discusses aspects of Kant's work on awareness of self.

Frege was the first to take the 'linguistic turn' so characteristic of twentieth century analytical philosophy; Richard DeVidi shows how this new approach to philosophical problems invests the connection between the linguistic act of referring to oneself and awareness of self with considerable significance. Frege's first faltering steps toward an account of first person reference are generally conceded to be unsuccessful; but the terms in which he sets the problem have been enormously influential on almost all subsequent work in the area. DeVidi argues, against Frege's critics and his revisionist friends, that Frege's approach to the theory of meaning requires only a very minor extension to deal with the peculiar features of first person reference quite successfully.

We did not commission a separate treatment of Wittgenstein's contributions to our topics for the following reasons. Wittgenstein's remarks fall mainly into two groups, the middle period remarks of the *Blue and Brown Books* (1933-4) and remarks from his final period in works such *Philosophical Investigations* (1953)*, Zettel,* and *Remarks on the Philosophy of Psychology*, though there also some relevant but highly compressed remarks in the *Tractatus* (1921). Shoemaker's and Evans' works, both of which had to be part of the volume in any case, discuss the most interesting of Wittgenstein's middle period remarks so it was not necessary to commission a separate study of them. Wittgenstein's final period remarks on the use of 'I' and related matters, many suggest, deny that when we use 'I', etc., in the relevant ways, we are using them referentially at all. Peacocke discusses these notions near the end of his paper. The absence of a separate study of Wittgenstein's final thoughts on self-awareness may come as a surprise. It reflects the following considerations. First, it is singularly difficult to get a clear, uncontroversial account of what Wittgenstein was trying to show us in his late remarks on first person pronouns, etc. If anything, it is even harder to see what the considerations were that, in his view, supported his remarks. Second, his remarks have persuaded almost no one who is not already strongly drawn to his general approach. For example, not a single contributor to this volume holds that the relevant uses of 'I', etc., are nonreferential and a number of them explicitly deny the claim. In short, the 'non-cognitivist' approach has had little influence on the work collected in this

volume.

Section II contains four papers that have achieved the status of classics in work on our topics. The paper that launched the whole recent dialogue was Hector-Neri Castañeda's famous work, '"He": A study in the logic of self-consciousness'. In this and other papers on the same topics, Castañeda argues that certain uses of 'I' and cognates are ineliminable. In particular, no description, not even one containing (other) indexicals, can be substituted for these uses. This is one of the two great claims that launched the recent discussion.

The second was first articulated by Sydney Shoemaker in 'Self-reference and self-awareness', the idea of immunity to error through misidentification of another as oneself. This immunity goes with another central peculiarity of self-awareness, what Shoemaker calls self-reference without identification. We introduced both ideas earlier.

In 'The essential indexical', John Perry picks up Castañeda's claim that nothing can be substituted for certain uses of 'I' and cognates and articulates it as the idea of the essential indexical. Building his account on what he needs to know in order to find out that a person making a mess in a store is him, Perry argues that without knowing that some part of a description applies to himself, without therefore knowledge of himself (and *as* himself), even a complete description of the person who is in fact making the mess need not be enough for him to find out that it is himself.

The chapter from Gareth Evans' 1982 book, 'Self-identification', is a massive, sprawling meditation on many aspects of self-reference and self-awareness. One suspects that Evans might have done some more work on it had he lived long enough. At any rate, it resists summarization. Shoemaker's two themes, self-reference without identification and immunity to error through misidentification, and their relationship to one another play a central role in the chapter. So does the idea that one must have more indexicals than 'I' and its cognates available. In particular, one must be able to locate oneself in the objective world (and so be able to use 'here') and know when a given time has arrived (and so be able to use 'now').

In Section III, we turn to more recent work on self-reference and self-awareness. All the

papers were prepared for this volume, though one of them, Peacocke's, was co-published elsewhere (Peacocke 1997) and Millikan's paper is a revision of her well-known 1990 critique of Perry. One of the first things to notice about these papers is how pervasively they have been influenced by the four classical works of Section II. Every paper of Section III criticizes papers in Section II. Indeed, every one of the former papers makes reference to at least two of the latter papers. One Section III paper, Ezcurdia's, also criticizes a Section III paper, namely, Millikan's. Some might suspect that the original authors could handle many of the criticisms laid out in Section III but all five papers raise significant issues and back their claims with considerable force of argument. In fact, it would extraordinarily interesting to set the Section III criticisms against the Section II claims that they aim to criticize and adjudicate the disagreements.

Ruth Millikan's criticism in 'The myth of mental indexicals' can be summarized quite succinctly: there is something in reference to self that is essential but it is not anything indexical. Why? The relevant uses of 'I' do not have features central of indexicals. For example, indexicals have no constant reference but 'I' as used by a given person always does. Moreover, any element that *is* indexical in uses of 'I' does not yield the information crucial to predicting what the user of 'I' will go on to do when, a central element in reference to self. And so on. In short, the element in reference to self that is essential is not an indexical.

In 'Thinking about myself', Maite Ezcurdia criticizes two authors who, she says, argue that there need not be any presentation of self when one refers to oneself using 'I' and cognates. Mellor is one. Millikan, the Millikan of 1990 and this volume, is the other. Mellor holds that the job that a representation of self is supposed to do is in fact done by context. If that were so, Ezcurdia argues, we would have no way to discriminate between an indexical referring to me and an indexical referring to the place I'm at. Since we can discriminate these things, Mellor must be wrong. In Ezcurdia's view, Millikan's emphasis on the relationship between an indexical and a referent specified by the semantic rules for the kind of indexical in question goes with a 'Millian' view of indexical reference, in which the semantics of an indexical are exhausted by its naming function and mode of presentation can play no role. Ezcurdia argues that this view leaves

Millikan with no way to distinguish between, for example, a use of 'I' presenting an object that happens to be myself and a use of 'I' presenting what I know to be myself.

In 'Introspective misidentification: an I for an I', Melinda Hogan and Raymond Martin go after the notion of immunity to error through misidentification. They do so primarily by presenting four putative counter-examples to the view. The counter-examples are all of the form: I am aware of someone having a feeling in the way that I am aware of feelings that I am having but it may in fact not be me who is having them. They consider objections to their counter-examples and reply to the objections. Two of the counter-examples are enmeshed in current controversies about identity of persons over time.

In 'First-person reference, representational independence, and self-knowledge', Christopher Peacocke separates a special kind of awareness of self in which one knows something about oneself, e.g., that one is seeing something, when this could not be inferred from anything about how the object of this act is represented. Peacocke calls these *representationally independent* uses of the first person. (His distinction is subtle and much more interesting than this gloss.) What interests him is that often representationally independent uses of 'I' and cognates give one knowledge, knowledge, for example, that it is oneself who is seeing something. He offers what he calls a *delta account* to explain this, according to which a subject *having* a property is sufficient in the relevant cases for a subject to *refer to itself*, and specifically, to *ascribe* the property to itself. He then applies the account to the illusion of a "transcendent subject" that he claims to detect in Kant, Schopenhauer and the early Wittgenstein, perhaps without recognizing that his delta account itself has strongly Kantian elements (see Brook, this volume). He concludes with a critical look at 'no-ownership' theories and theories that the relative uses of 'I' and cognates are not referential.

Seager's 'The constructed and the secret self' can be seen as a step toward meeting a need that Evans articulates (1982, 205; this volume, xxx). Says Evans, we cannot give a complete account of self-reference and self-awareness without a complete theory of mind, though we can make some progress while we wait. Seager's paper takes some steps toward

developing such as general theory of mind. He urges that from the perspective of a representational theory of mind, it is natural to arrive at a dual view of selfhood, i.e., of whatever it is that realizes personhood and carries sameness of personhood across time. He then applies the resulting duality to a wide range of phenomena, all the way from Williams' puzzle cases about personal identity to such mundane matters as the relative cognitive impenetrability of illusions such as the Müller-Lyer arrowhead illusion and the surprisingly tenacious conflicts between, e.g., our considered values and what we desire that we find in ourselves from time to time, most of us at any rate.

Rich in precursors and stimulated by Castañeda's and Shoemaker's seminal papers, the work of the past thirty-five years on self-reference and self-awareness has generated a wealth of deep, sophisticated philosophy. We hope that this volume conveys something of that richness.

Andrew Brook

Richard DeVidi

Carleton University, Ottawa, Canada