

THE VARIETIES OF SELF-KNOWLEDGE

A Thesis

by

ANTON SERGEEVICH KABESHKIN

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

MASTER OF ARTS

May 2011

Major Subject: Philosophy

THE VARIETIES OF SELF-KNOWLEDGE

A Thesis

by

ANTON SERGEEVICH KABESHKIN

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

MASTER OF ARTS

Approved by:

Chair of Committee,
Committee Members,

Head of Department,

Robert W. Burch
Benjamin McMyler
Steven Smith
Daniel Conway

May 2011

Major Subject: Philosophy

ABSTRACT

The Varieties of Self-Knowledge. (May 2011)

Anton Sergeevich Kabeshkin, B.S., South Ural State University

Chair of Advisory Committee: Dr. Robert W. Burch

In this thesis I consider the problem of the distinctiveness of knowledge of our own mental states and attitudes. I consider four influential approaches to this problem: the epistemic approach, the “no reasons view,” the neo-expressivist approach and the rational agency approach. I argue that all of them face serious problems. I further argue that many of these problems are connected with the lack of fine-grained enough classification of the entities with respect to which we have self-knowledge. I suggest such a classification, distinguishing passive occurrent mental states, mental actions and standing attitudes, and argue that we should treat each of these categories separately for the purpose of explaining self-knowledge of them. I discuss in detail self-knowledge we have with respect to two of these categories: standing attitudes and mental actions. On my account self-knowledge of standing attitudes stands in a derivative relation to self-knowledge of other kinds. In my discussion of self-knowledge of mental actions I establish that we have a distinctive non-observational kind of self-knowledge and show some specific characteristics of this kind of self-knowledge. In the end I attempt to relate self-knowledge of mental actions to practical knowledge in the ordinary sense of skill.

ACKNOWLEDGEMENTS

I owe special thanks to the chair of my committee, Robert W. Burch. The discussions with him were absolutely essential for crystallizing the ideas which are developed in this thesis. His criticism, advice and corrections were always helpful for my work.

Thanks to the other members of my committee, Benjamin McMyler and Steven Smith, for their advice with literature on the early stage of the project. Dr. McMyler's comments on the drafts of various parts of this work and keen criticism always gave me impulse to spell out my ideas in more detail.

Great thanks to all of the students, faculty and staff in the philosophy department of Texas A&M University for their helpfulness and for creating the ideal climate for doing philosophy.

TABLE OF CONTENTS

	Page
ABSTRACT	iii
ACKNOWLEDGEMENTS	iv
TABLE OF CONTENTS	v
CHAPTER	
I INTRODUCTION.....	1
II THE EXISTING APPROACHES AND THEIR CRITICISM	6
The Exposition of the Existing Approaches.....	6
Criticism of the Existing Approaches	18
III DIFFERENT KINDS OF SELF-KNOWLEDGE.....	33
Classification of the Objects of Self-Knowledge.....	33
Self-Knowledge of Our Attitudes	37
IV SELF-KNOWLEDGE OF OUR OWN ACTIONS	41
Agency and Self-Knowledge: Some History	41
Overt Bodily Actions and Self-Awareness	45
Mental Actions: Some Preliminaries.....	48
Criticism of the Rival Accounts	53
Self-Knowledge of Mental Actions: More Details	60
V CONCLUSION	69
REFERENCES.....	71
VITA	76

CHAPTER I

INTRODUCTION

Among things we know and are able to speak about there is a certain class of things such that we assume that a competent speaker has a special authority with respect to knowledge about them. This class includes certain facts about us. For example, I am usually the best person to answer such questions as whether I am in pain now, or whether I prefer to eat some fish rather than meat, or whether I fear spiders. Of course, there are some cases when for some reason I am not aware of my wish or fear and someone else can judge about it better on the basis of observations of my behavior. However, as a rule, such judgments are overridden by my own resolute claims to the contrary, if my sincerity is granted. And it is much more difficult to find cases in which I mistakenly judge that I am in pain.

Of course, not all facts about us have this kind of privilege. Facts about my appearance, for example, can be known by anyone with the same certainty as by me. The facts which I typically know with greater certainty than anyone else include only those which I know “from the inside” or “directly.” These are, in other words, facts about my mental states, although what exactly counts as a mental state and what is the nature of mental states is debatable.

The problem of self-knowledge, that is, knowledge of one’s own mental states and acts as distinctive in comparison with knowledge of other kinds, has drawn

This thesis follows *The Chicago Manual of Style*.

considerable attention in recent decades.¹ Yet there is no general agreement even concerning the formulation of the principal question to be answered. Some authors are concerned with the problem of how we know our mental states and why this kind of knowledge is more certain than any other kind of knowledge. Others speak about the distinctive first-person authority (or privilege) of self-ascriptions of mental states. Later in this work I will argue that such a shift of attention towards the problem of the authority of self-ascriptions substitute the communicative problem — that is, the question Why do we generally trust people’s utterances about their mental states? — for the problem of knowledge of mental states. Such substitution may bring misleading results, for the problem of knowledge of mental states is more fundamental than, and should be solved independently from, the subordinate communicative question.

Another source of problems is the question whether a single explanation of self-knowledge about all mental states should be given, or rather multiple explanations for different kinds of such states are needed. Those authors who are inclined to the second option typically distinguish two kinds of self-knowledge.² I will argue that different kinds of self-knowledge should indeed be recognized and I will even radicalize the existing approaches in this direction, distinguishing at least three different categories of the entities with respect to which we have distinctive self-knowledge (allowing also for intermediate cases). I will argue that knowledge of each category needs a separate

¹ Some essays are collected, for instance, in Smith, Wright and Macdonald 1998, and Cassam 1994. Other references will be given in the course of the paper.

² Moran 2001, p. XXXIII distinguishes occurrent mental states such as sensations and feelings, and standing attitudes (beliefs, intentions and so on). Boyle 2009 distinguishes (following Kant) active and passive kinds of self-knowledge, the first based on spontaneity of an agent, the second dependent upon some other mechanism, such as “inner sense”.

explanation, although these explanations need not be unrelated. In particular, we should distinguish occurrent mental acts (such as making an assumption, silent utterance or imagining something), passive occurrent mental states (sensations, feelings), and dispositional attitudes (non-occurrent beliefs, desires, etc., although there are more subtle distinctions within the last category, which I will briefly touch).³ I will further argue that many problems of the existing accounts of self-knowledge are a result of the lack of an adequate classification like the one suggested above. In this work I will also consider the main accounts suggested specifically for explaining self-knowledge of our actions in general and of our mental actions in particular on the one hand, and self-knowledge of our beliefs on the other. I will argue that the most promising of these accounts are those that relate our self-knowledge to our agency (primarily in the case of actions and in a derivative way in the case of beliefs).⁴

In Chapter II I will provide an outline of the main accounts of self-knowledge suggested in the literature and then will criticize them. In the first section of that chapter, I will consider the epistemic approach (or perceptual model), the “no reasons view,” the neo-expressivist account of self-knowledge, and the approach that links self-knowledge and agency. In the second section of Chapter II I will criticize these accounts and will show that none of them is adequate for the whole range of phenomena with respect to which we have self-knowledge. In Chapter III I will argue that any adequate theory of

³ It should be noted that I use the expression “dispositional attitudes” in this paper quite broadly. Roughly it means just “that, which belongs to our mental life but is not occurrent,” while an “occurrent state” is that which consists in subject’s possession of some categorial mental property. There is, for example, an important difference between “powers” or capacities and just reliable dispositions but, so far as I can see, my arguments do not hinge on this distinction. In the paper I will sometimes use just “attitudes” for the sake of brevity.

⁴ For the sake of brevity I will often call this account “the agency approach” in the rest of the work.

self-knowledge should recognize different categories of entities with respect to which we have self-knowledge. Moreover, one should not start from the assumption that self-knowledge of all these entities can be explained in a uniform way. In the first section of Chapter III I will present and argue for my classification of entities with respect to which we have distinctive self-knowledge. In the second section I will attempt to explain failures of the version of agency approach which links self-knowledge of our beliefs and intentions with rational deliberation. After that I will suggest an alternative account of self-knowledge of our beliefs and intentions based on my classification developed in the first section.

In Chapter IV I will argue that self-knowledge of our own actions in general and of our mental actions in particular is based upon non-observational knowledge of a special kind. In the first section I will provide a historical overview of this position. In the second section I will raise some objections to the idea of distinct non-observational knowledge of our own actions and will outline some alternatives to this idea. However, these alternatives work only with respect to our overt bodily actions and not with respect to our mental actions. Thus, in the rest of Chapter IV I will discuss self-knowledge of mental actions. In the third section I will discuss what counts as a mental action. In the fourth section I will present and criticize the accounts of self-knowledge of mental actions which are not based on the recognition of a distinct source of knowledge. This criticism, I will suggest, opens the way for the acceptance of the idea that self-knowledge of our mental actions is non-observational, non-inferential and yet rationally justified. Finally, in the fifth section I will attempt to elaborate this idea and to provide

more substance to it. In particular, drawing from the works of G.E.M. Anscombe⁵ and Kieran Setiya,⁶ I will attempt to link it with the so-called “knowledge how.” However, this suggested link will have only the hypothetical status, and I will raise some problems for this development. In the conclusion I will briefly summarize the argument of the work and will point out to the connection between self-knowledge of our attitudes and self-knowledge of our actions in light of my treatment of self-knowledge of our actions in Chapter III.

⁵ Anscombe 2000.

⁶ Setiya 2008 and 2009.

CHAPTER II

THE EXISTING APPROACHES AND THEIR CRITICISM

The Exposition of the Existing Approaches

In this section I will expound in turn four prominent accounts of self-knowledge to provide the reader with an overview of the field. I will postpone the criticism of these accounts until the next section except insofar the criticism of rival approaches is necessary to elucidate the motivation for the development of other approaches.

The Epistemic Approach

There are multiple versions of the epistemic approach, but they all concur that we know about our mental states because we somehow “detect” them via some epistemic mechanism which works on the personal level. The special certainty of this knowledge is, accordingly, accounted for by the high reliability of this epistemic mechanism. Finkelstein (2003) calls the approach in general “detectivism” and distinguishes “old detectivism” from “new detectivism”. The old version is connected with what is often called the “Cartesian paradigm.” It considers both our mental states and our attitudes as special objects that belong to consciousness as opposed to the physical world, and that should be perceived in order to be known. It then posits that we possess some very special “mechanism” called introspection, by means of which we can indeed perceive these states and attitudes. In this respect it is similar to our ordinary senses but, unlike

them, it is infallible and immediate. Through it we can perceive our states without a possibility of mistake. These properties, of course, account for the first-person privilege and the presumption of truth of self-ascriptions of mental states. However, because it can hardly be the case that any bodily process can be infallible, the infallible introspection can only be a capacity of an immaterial soul. Thus, this view naturally leads to dualism.

Because of this association of “old detectivism” with dualism, old detectivism is not considered to be a viable option by most of the contemporary philosophers. Thus, those who are still attracted by the epistemic approach usually hold that we know our mental states via a fallible perceptual capacity (“new detectivism” in Finkelstein’s terminology). One way to develop this view is to posit that we have an introspective ability that is more akin to ordinary sense perception than the Cartesian introspection, and so it can be mistaken just as our eyesight or hearing. This is the version of epistemic approach I will be mainly concerned with throughout the work. One of its prominent contemporary proponents is D. M. Armstrong.

According to Armstrong,⁷ introspection has a structure similar to that of perception. In particular, in introspection we should distinguish that which introspected from the introspecting, as in perception we distinguish the perceived object and the act of perception. Unlike in perception, however, in introspection both the entity introspected and the introspecting are mental states; but they are distinct mental states. Their distinctness is one of Armstrong’s primary grounds for asserting that introspection is fallible (I will consider his arguments for this shortly). Moreover, not only can

⁷ Armstrong 1963 and 1994.

introspection be mistaken, but it also does not perceive all the mental events. Thus, Armstrong denies what is sometimes called “omniscience” in discussions of first-person knowledge, that is, the claim that all mental events are necessarily conscious.⁸ Indeed, this is quite reasonable provided that the mental state of perceiving is distinct from the mental state perceived, which Armstrong maintains.

Armstrong not only explicitly asserts the fallibility of introspection, but also argues for it. To be more precise, he argues that introspection cannot be *logically* infallible. One of his arguments is a familiar Wittgensteinian one that in cases where mistakes are not possible it is inappropriate to talk about knowledge as well: “We can speak of gaining knowledge only in cases where it makes sense to speak of thinking wrongly that we have gained knowledge.”⁹ This argument, however, does not seem to be conclusive. One can contend that a true proposition that one holds to be true for good reasons must count as knowledge whether or not it is possible for a subject not to hold it to be true. And the fact that a proposition cannot be false is arguably quite a good reason to hold it true.

Another argument is based on the distinctness of the target mental state of introspection and the state of introspecting itself. The point is a Humean one that it is always possible to conceive one of the “distinct existences” without the other. But, the argument goes, what is conceivable is logically possible, and thus, in our case, it is logically possible that the introspecting mental state occurs while the mental state introspected does not (or vice versa). Note that this argument rules out only logical

⁸ See, for example, Alston 1971, 228.

⁹ Armstrong 1963, 422.

infallibility of introspection. Armstrong does not consider whether introspection can be nomologically infallible, that is, always yield correct results because of certain psychological or, rather, physiological (since Armstrong identifies all mental states with brain states) laws.

In any case, we can see that one of the most important characteristics of the epistemic approach is the distinctness of the entities perceived by introspection and the states of perceiving themselves. This assumption is rejected by at least some of the versions of the view which will be considered in the next subsection.

The “No Reasons View”

As we have seen, the epistemic approach involves the view that we possess a capacity to perceive our own mental states. For its adherents our beliefs about our own mental states are justified by deliverances of this capacity in a way parallel to how our beliefs about the external world are justified by deliverances of ordinary perception. Some authors which are not content with the epistemic approach have argued that there is no such a capacity. Rather, they maintain, our second-order beliefs (that is, beliefs about other beliefs) are given rise to “automatically,” without mediation by personal-level reason-giving capacity or mechanism. Because of the absence of such a capacity we are not able to provide reasons for having second-order beliefs: we simply do not have such reasons.

At this point the “no reasons view,” as it is called by some authors, is supported by our actual linguistic behavior. When we make reports about, say, our current sensations or beliefs, we do not provide further reasons for our knowledge that we have

those sensations or beliefs. In the unlikely situation when someone will ask the reporting subject to justify her claim that she has those sensations or beliefs, the subject will not respond that she has perceived them. Rather, she will wonder if the questioner has understood her well enough.

Thus, partisans of the “no reasons view” accept the apparent lack of reasons for the beliefs about our own mental states at face value. Still, they want to maintain that these beliefs are, at least as a rule, true. So they need to show how the correspondence of the presence of first-order mental states, acts or attitudes and of the possession of beliefs about them is ensured. Now, there are different ways to do this. One way is simply to postulate the existence of a subpersonal causal mechanism that reliably produces beliefs about the presence of our mental states in case when these mental states are present. Since this causal mechanism operates only at the subpersonal level, it is *merely* reliable in a sense that it does not provide epistemic reasons for judgments and beliefs to the effect that target mental states are present, which a subject can provide for the justification of these judgments and beliefs. Such view is naturally accompanied by the reliabilist epistemology.¹⁰

There are various reasons for dissatisfaction with this reliabilist version of the “no reasons view.” An alternative to this version is what David Finkelstein called the constitutive approach. According to such an approach the lower-order and the higher-order mental states are not distinct entities. Thus, this version of the “no reasons view” rejects not only the core of the epistemic approach, namely the existence of a personal-

¹⁰ In this passage I am indebted to Peacocke 1998, 76-7.

level capacity to “detect” our own mental states, but even the assumption on which the epistemic approach is founded, that is, the distinct existences of lower-order mental states and judgments or beliefs about those states.

This view as applied to entities such as intentions is sometimes supported by reference to Wittgenstein’s criticism of certain forms of Platonism about intentions. This line of argumentation suggests that subsequent judgments about the content of prior intentions somehow determine the content of those intentions. Thus Crispin Wright: “The role of subsequent judgement is indeed not to mediate somehow in the connection between the content of an intention and its execution... but rather to enter into determination of *what* the content of an anterior intention is to be understood as having been.”¹¹ And, a little further in the text:

The Platonist mythology is a mythology of such constitutive independence... Against this Platonism I want to set what I take to be an idea of Wittgensteinian authorship, although it is of course very familiar from the writings of Davidson: that the content of a subject’s intentional states is not something which may merely be accessed, as it were indirectly, by interpretative methods... but is something which is intrinsically sensitive to the deliverances of best interpretative methodology. That is a methodology which in principle must include within its conspectus the whole sweep of a subject’s sayings and doings, including future ones, without bound.

Whether or not this is a fair interpretation of Wittgenstein or Davidson is a controversial question (McDowell, for instance, resists Wright’s reading of them). What is important for us here, though, is that such an interpretation provides some flesh to the constitutivist view that claims higher-order judgments to be constitutive of lower-order mental states. Of course, while putatively Wittgensteinian or

¹¹ Wright 1998, 29.

Davidsonian arguments might at least make it plausible that higher-order judgments are constitutive of our intentions, this can hardly be said about other kinds of mental entities. But, as I have said earlier, I will reserve the criticism for the next section.

Neo-Expressivism

Another interesting approach that purports to explain distinctive first-person knowledge is neo-expressivism. This view shifts the focus of attention towards the way we talk about our own mental states or otherwise express them, and thus attempts to shed light on our relation to these states via the analysis of our expressive capacities.

Neo-expressivists claim that what is distinctive in our relation to our mental states is that our self-ascriptions of these states can normally directly “express” these states.¹² By “expression,” for example, Bar-On means “giving vent, airing” and so on, that is, manifesting one’s mental state by either action or self-ascription. Thus, one can express her pain by screaming, by making a grimace or by some kind of utterance. Non-verbal expressions are sometimes called “natural expressions.” In the case of verbal expression an expressive utterance can be either self-ascriptive, such as “I feel pain,” or not self-ascriptive (for instance “It burns!”). A self-ascriptive utterance is called an avowal. For the present goals the most important difference between avowals and other kinds of expressive behavior is that avowals, unlike natural expressions, also have

¹² Recent publications of the proponents of this view include Bar-On and Long 2001, Finkelstein 2003 and Bar-On 2004.

semantic content that self-ascribes the same mental state as the one they express.¹³ To explicate how this is possible, Bar-On distinguishes two aspects of avowals: avowals as acts and avowals as products.¹⁴ Avowals as acts express relevant mental states, while avowals as products ascribe the same states to a speaker. For example, on the neo-expressivist view an utterance “I feel pain” expresses (shares, gives vent to, airs) the speaker’s pain in virtue of its being a specific kind of act that has a similarity to the more simple means of expression. But, at the same time, this utterance ascribes this same pain to the speaker in virtue of its linguistic meaning.

Now, on the neo-expressivist account, avowals enjoy the distinctive first-person authority in the same manner and for the same reasons as do natural expressions of mental states such as screams, twitches, smiles. More specifically, avowals inherit their distinctive first-person authority from natural expressions. As Bar-On argues,¹⁵ children learn to substitute avowals for natural expressions in the course of their development and start using avowals to express their mental states. Thus, when someone uses an avowal to express her own mental state, she does not base the self-ascription which is contained in the avowal on some prior epistemic judgment to the effect that she is in that state, as one does not base her smile or scream on a prior epistemic judgment. Consider this illustration from Bar-On (which also shows her view on the substitution of avowals for natural expressions in the course of child’s development):

[A] small child, Jenny, eagerly reaching for a teddy bear. Jenny simply wants the toy, and her reaching for it directly reveals her desire, quite

¹³ Verbal expressive utterances that are not avowals, of course, also have semantic content, but it does not self-ascribe states expressed to speakers.

¹⁴ Bar-On 2004, 251 ff.

¹⁵ Ibid., 286 ff.

independently of any judgment on her part to the effect that she wants the teddy. Using previous jargon, we might say that there is no epistemic distance between Jenny's behavior and her desire for the toy. Her present state is in no way an epistemic target for her; she simply gives vent to it. And in normal circumstances, her audience will be directly responsive to the state they perceive her to be in; barring contravening reasons, they will simply hand her the toy.

But now consider another episode in which Jenny emits a certain sound ("Uh!"), or calls out "Teddy!", as she reaches for the toy. And finally, consider an episode in which she avows "I want Teddy," perhaps with no reaching at all. Intuitively, the verbal emissions, just like the reaching behavior, which we would consider a natural expression of Jenny's desire, may all 'come directly from' the child's desire. They seem equally 'pressed out' from her, and they appear no more driven by a prior deliberation, consideration, or determination regarding how things are presently with her. The verbal utterances seem to be equally expressive of her desire for the teddy.¹⁶

The fact that avowals are not grounded on prior epistemic judgments is what explains avowals' epistemic security. Even if one can make a mistake in avowal (and Bar-On leaves open such a possibility), such a mistake will not be akin to the types of mistakes associated with judgments about the external world. Rather, such mistakes will be connected with failures of rationality (as in self-deception or wishful thinking) or failures of expression (as when one attempts to express pain when one does not actually feels it but merely anticipates it¹⁷).

Thus, neo-expressivism ties our distinctive authority over our own mental states to our ability to express them and then analyses the kinds of self-expressions and their structure. Ultimately, on this account the peculiarity of self-ascriptive expressions is what explains our first-person authority.

¹⁶ Ibid., 242.

¹⁷ Ibid., 322.

The Rational Agency Approach

Another promising account of self-knowledge links our distinctive first-person authority to our capacities as rational agents. In this subsection I will consider this account as it is developed by Richard Moran in his book *Authority and Estrangement* (2001). Although there are also developments in this vein by Matthew Boyle and especially by Sebastian Rödl, I will here consider only Moran's view, for while I think it is on the right way, it illustrates one important mistake which can be made on this route.

In his account Moran first distinguishes self-knowledge of standing attitudes such as beliefs and intentions from self-knowledge of passive occurrent mental states. He writes that it is plausible that two different stories should be told about these two kinds of self-knowledge. Then he explicitly narrows his investigation to the case of standing attitudes.¹⁸ His basic idea is that we know our attitudes not by observation, not by any kind of "inner sense" (thus Moran rejects the epistemic approach), but by actively forming them, by making up our mind. To appreciate this idea, it is instructive to compare two different stances which can be taken with respect to our attitudes: a practical or deliberative stance and a theoretical stance.¹⁹ In taking the theoretical stance one asks oneself whether, for example, she wants to do A, whether she believes that p, etc., and answers this question by ascertaining empirical evidence about oneself. This evidence can be gathered in more or less the same way one can gather facts about someone other than oneself: by observing one's overt behavior, reflecting on one's biographical facts, using methods of cognitive science or neurophysiology, etc. Perhaps,

¹⁸ Moran 2001, 9-10.

¹⁹ Ibid., 55-60.

unlike in the case with other people, one can also use introspection to figure out whether one wants to do A, believes that p, etc., but the question that can be answered in this way is still a question asked in a theoretical stance. By contrast, questions asked in a deliberative stance are related to the process or to the outcome of making up our minds. Such questions as “do I intend to do something?” or “do I believe that...” asked in this sense will be answered not by the psychological or behavioral observation of the person who asks them, but by forming an intention or by considering evidence for and against the belief in question, respectively. So, this kind of deliberative reflection involves reasoning, theoretical in the case of belief²⁰ and practical in the case of intention. Moran notes that this reasoning is something more than merely normative appraisal of our attitudes or matters of fact, arguing that normative appraisal may well be applied to the past attitudes as well as to the present, but we cannot form our past attitudes. Deliberation ends not with a normative judgment but with an intention to do something, a desire for something, a belief, etc.—that is, with the attitude in question itself.

The deliberative stance is characterized by the so-called transparency condition.²¹ Consider the case of belief. When deciding what to believe, our gaze is, so to say, directed to the world. In order to answer the question whether to believe that p, we are considering reasons for the entirely different question: whether p is the case. While considering the similar question about some other person, we do not conflate the psychological question whether that person believes that p and the world-directed

²⁰ Here it is important not to confuse theoretical reasoning involved in the belief formation with theoretical stance, which we can take with respect to our attitudes, as described earlier.

²¹ Ibid. 60-65.

question whether p. The same distinction clearly can be applied in our own case as well, for we are as susceptible to error as anyone else. Nevertheless it would be paradoxical to say both “I believe that p” and “it is not the case that p” (this is so-called Moore’s paradox). It is worth noticing, though, that even in our own case we *can* sometimes ask questions about our beliefs in the “theoretical” stance when we will be looking precisely for the evidence of the presence of the psychological state, such as certain behavior. This might happen in a context of psychoanalytic treatment, when the belief in question is unconscious, or of analysis of one’s feelings which is characteristic for heroes in the literature of romantic love.

Although his paradigm case is the case of beliefs, Moran thinks that the same kind of transparency is also applicable to attitudes other than beliefs, such as fears, desires, intentions, even emotions.²² He is trying to find for them analogues of the transparency to the world and argues that, for instance, fears or desires can be formed and in the fundamental case are formed via weighing reasons for and against fearing something or desiring something, with a gaze directed to the world. For instance, I can consider whether something is threatening and thus whether I am to fear it.

These considerations still leave the main question open, namely the question how exactly the unique first-person capacity to be a rational agent and to form one’s attitude via deliberation explains first-person *knowledge*. For one can grant that we can deliberate and that as a result of deliberation a certain attitude will be formed in me, but argue that it is not evident that I will thereby know that I have this attitude. I am not sure

²² See, especially, Moran 2002, 206 ff.

that Moran provides an explicit and satisfactory answer to this challenge. However, I think that this gap in his account can be plausibly filled in.²³ It could be noticed that an action explanation, that is, an explanation why I am performing an action, is the same thought with which I conclude my practical deliberation whether to do that action. The same goes for belief explanations and conclusions of theoretical deliberations. For instance, if I have proved a logical theorem, the proof I have made will also be an explanation why I hold that this theorem is valid. Because of this, in the (successful) end of theoretical reasoning I give an answer which is also a reason for my holding the belief that resulted from this reasoning (and likewise for practical reasoning and intention). Furthermore, it seems plausible to assume that if I know *why* I hold some belief, I also know *that* I hold this belief. Thus, if I went through the deliberation which issued in a belief, I know the reasons for which I came to have that belief, hence I know why I now hold it, and hence I know that I hold it. The same reasoning can be *mutatis mutandis* applied to intentional actions.

So, on this approach, our capacities for deliberation explain the first-person authority at least with respect to our beliefs and intentions. How far the rational agency approach can be extended is an open question.

Criticism of the Existing Approaches

In this section I will provide the criticism of the approaches outlined above. As far as the

²³ In what follows in this paragraph I am much obliged to Sebastian Rödl's recent book *Self-Consciousness*, in particular to chapters 2 and 3. Rödl's book is very rich and intricate, and I will only use one of his ideas, and even that only in a sketchy way.

epistemic approach, the “no reasons view,” and the neo-expressivist approach go, I will criticize them while taking into account that they are supposed to explain all kinds of self-knowledge, since their proponents typically do not restrict the application of these approaches to a particular class of entities. On the other hand, since proponents of the rational agency approach do restrict it to standing attitudes and dispositions such as beliefs and intentions (with a possible expansion to the attitudes such as desire or fear), I will take into account this self-restriction.

Criticism of the Epistemic Approach

I have already mentioned that the old version of the epistemic approach leads to the Cartesian dualism, which makes it rather unpopular among contemporary philosophers. Nevertheless, part of my criticism will be directed against all kinds of the epistemic approach. First, however, I will expound several objections against specifically the “new epistemic approach,” as they are developed by David Finkelstein.²⁴

First, it is important to notice that first-person authority is only applicable to conscious mental states and attitudes. If one is angry at someone but is not conscious of this anger, one’s self-ascriptions of it will be based on the sources of information similar to those we use to judge about other people’s attitudes. Hence such self-ascriptions are no more authoritative than judgments of others who can observe the subject’s behavior. According to the epistemic approach what makes attitudes conscious and explains the first-person authority we have with respect to them is some form of the inner sense.

²⁴ Finkelstein 2003, 9-27.

Similarly, if there are unconscious sensations (and we have seen that, for example, Armstrong holds that there are), only those sensations are conscious which are perceived by the inner sense. Now, if this inner sense is similar to the ordinary “outer” senses, it is not clear why its objects should be conscious. That is to say, the attempt to model the inner sense on the outer senses fails at this point: the inner sense should significantly differ from the outer senses in that it renders what it perceives conscious. And merely postulating that it does so does not give us an explanation why it does. This point undermines the aspiration of the proponents of the new epistemic approach to make a break with the Cartesian tradition of construing the inner sense as something very special.

Furthermore, whereas some mental entities do have a certain phenomenal character, such as that of pain, anger and, probably, fear or desire, there are others that lack it. It is not clear whether intentions have it or, if they do, whether it differs from that of desires. But it seems clear that beliefs do not have a specific phenomenal character: there seems to be no “feeling like” having a belief. Thus, it seems that the inner sense model does not accommodate all the mental phenomena with respect to which we have characteristic self-knowledge.

Now, the epistemic approach faces problems that are not exclusive for its modern versions.²⁵ If indeed we perceive our mental states, acts and attitudes by the inner sense, the question arises: How do we know that these mental states, acts or attitudes are our states rather than someone else’s? If we perceive just a mental state, it is not clear why

²⁵ The argument is adapted from that of Shoemaker’s (1998, 128).

should we ascribe it to ourselves. A possible answer could be that one does not perceive just a mental state but oneself possessing that mental state. But this answer only pushes the problem a step back, for the new question arises: How is this self identified as one's own self? It does not seem that we ever do identify some self as our own. But even if the proponent of the epistemic approach will attempt to disregard the appearances and assert that we do identify our selves given in introspection as ours, say, as possessors of unique sets of properties which we know before the act of introspection, we can further ask: How do we know *those* properties from the supposedly unique sets themselves? If they are obtained by introspection, then infinite regress is threatening. If they are not obtained by introspection, then we possess some self-knowledge that is not provided by introspection, and then it is shown at the very least that introspection cannot explain all self-knowledge.

Shoemaker considers the following response to this argument: one can argue that only one self, namely my own, can possibly be the object of my inner sense, and thus it is not necessary to identify it. Shoemaker notices that "this amounts to saying that I can infallibly identify the observed self as myself by the fact that it is introspectively observed by me." However, this response presupposes that I know that it is indeed my self who is the subject of this particular introspection. But then *this* will be the piece of self-knowledge, which is not obtained via introspection, and the main point of the argument remains untouched. Thus, the epistemic approach cannot explain self-knowledge of all kinds, and even if there is introspection of some sort, its successful operation presupposes certain self-knowledge.

Criticism of the “No Reasons View”

As I have mentioned during the exposition of the “no reasons view,” some of its versions are dependent upon the reliabilist epistemology. Now there are many people who object to it, for on their view beliefs that are simply reliably caused by non-rational mechanisms are not justified. Reliabilism offers us, in John McDowell’s words, “exculpations where we wanted justifications.”²⁶ And if one rejects reliabilism in general, it is very odd to make an exception in one particular case and allow the application of the reliabilist epistemology to the case of self-knowledge only.

Moreover, although some mental entities with respect to which we have self-knowledge lack phenomenal quality, as was noticed above, most of them do have such quality. As Christopher Peacocke notes,²⁷ such mental states and attitudes can give reasons. Specifically they can provide “a conceptually equipped thinker” with reasons for self-ascription of these mental states and attitudes. And if a mental state gives reasons for a belief or judgment about this state, then they should be distinct states, since we do not count as justifications of belief that *p* references to this same belief. This argument works against the constitutive versions of the “no reasons view” as well as against the less sophisticated versions.

A further objection to the specifically constitutive accounts runs as follows. On these accounts mental states or attitudes and the corresponding higher-order beliefs are identical. Because of that, it seems that such accounts are committed to the view that

²⁶ McDowell 1994, 8. McDowell’s phrase at that place is directly aimed at what he calls “the Myth of the Given,” but he also applies it to any epistemology which offers merely external constraint for our beliefs, rather than justifications.

²⁷ Peacocke 1998, 83.

either there are no wrong higher-order beliefs (if higher-order beliefs constitute our possessing lower-order mental states or attitudes) or that there are no lower-order states or attitudes, such that there are no corresponding higher-order beliefs about them (if it is part of our possession of lower-order mental states or attitudes to have corresponding higher-order beliefs about them). However, there are counterexamples to both of these commitments.

The example of the wrong higher-order belief about one's attitude would be a higher-order belief obtained in an "indirect way," for example by observing one's own behavior. As I have noticed earlier, it is quite possible to form beliefs about one's own attitudes in this way. And of course it is possible to make a mistake when forming beliefs in this way. Thus, it is possible to have a mistaken higher-order belief about one's attitude, which seems to be precluded if having a higher-order belief is what constitutes possessing a lower-order attitude.²⁸ As for an example of a lower-order state or an attitude without a corresponding higher-order belief about it, one can just take any unconscious state or attitude. Even if one objects to the idea of unconsciousness in any sense reminiscent of psychoanalysis, I think it is clear that there are perceptions to which we just do not pay attention and so do not form beliefs about them. Incidentally, the evidence is collected that even quite complex mental phenomena such as daydreaming may remain undetected by a person engaged into daydreaming.²⁹

I conclude that the "no reasons view" does not work as an explanation of self-knowledge either.

²⁸ Martin 1998, 107.

²⁹ See, for example, Schooler 2002.

Criticism of Neo-Expressivism

In my criticism of the neo-expressivist approach I would not like to downplay some of its interesting insights in explaining the distinctive authority of avowals. However, I contend that we should distinguish two different questions which are not reducible to each other. One question is “Why do we normally presume that people’s self-ascriptions of mental states, if sincere, are true?” I suggest to call it the question of “the communicative aspect of self-knowledge.” It may well be that the neo-expressivist account provides a correct answer to this question. Another question is “How do we obtain knowledge of our own mental states and acts and why is it so reliable, if not infallible?” This is the more fundamental problem of self-knowledge. I contend that neo-expressivism does not resolve this problem.

To see this, we need to consider how neo-expressivism explains the difference between conscious and unconscious mental states or attitudes. This explanation is important because distinctively authoritative self-knowledge can be knowledge only of our *conscious* mental states, acts or attitudes, for the ways to get knowledge of unconscious entities are essentially the same we use to get knowledge of other people’s minds. For neo-expressivists the condition for a mental state’s or an attitude’s being conscious has the form “The mental state X is conscious if the subject who has it is able to avow it.”³⁰

Now it is plausible that there are counterexamples to this condition. Consider the case of a normal adult who has recently sustained damage, say, to his Broca’s and

³⁰ Cf. Finkelstein 2003, 120: “Someone’s mental state is conscious if he has an ability to express it merely by self-ascribing it.”

Wernicke's brain zones and is also paralyzed, say, as a result of a car accident. Let us stipulate that in this situation because of the brain damage this person is unable to express his mental states in speech or in thought; nor is he able to express them even by gestures or mimicry (because of the paralysis). Yet it seems plausible that he could be conscious of at least such mental states as pain or other sensations in just about the same way that normal adults are, since he underwent normal mental development, including linguistic acquisition, and since his psychic setup is thus as developed as anyone else's is. Perhaps he is not conscious of his beliefs (that is, he is not consciously believing something) or other attitudes, but there are no apparent reasons for denying that he is consciously experiencing phenomenal qualities of his occurrent mental states.

Some might think that this counterexample or similar counterexamples are inconclusive. Yet it seems that they at least make it plausible that consciousness of at least occurrent mental states and an ability to express them are different things and, moreover, one can exist apart from the other.³¹ I think that what brings neo-expressivists to the conclusion that they are connected more tightly is their failure to distinguish the kinds of self-knowledge we have with respect to attitudes and with respect to mental acts or passive occurrent mental states. Now I will suggest an argument that shows more formally than the counterexample above that consciousness of at least occurrent mental states is not to be explicated in terms of an ability to avow those states.

Let us accept the premises:

³¹ For the sake of simplicity in the rest of this section I will sometimes write "consciousness of" a state where Finkelstein would have "being consciously in" that state.

1. The ability to avow or otherwise express one's mental states, acts or attitudes is a disposition.

This is so because any ability or “power to do” is a disposition, broadly speaking (that is, not a categorial property or an occurrent state).

2. The property of a state's or an act's being conscious is a categorial property.

I will discuss this premise after the full exposition of the argument.

3. One entity cannot be both a disposition and a categorial property or an occurrent state.

From these premises it immediately follows that

4. Being conscious of a mental state and being able to avow or otherwise express it are not the same things.

This is just a preliminary result, and it does not on its own preclude the conditional analysis of consciousness of mental states, which Finkelstein provides. Even if being conscious of a state and being able to avow it are not the same thing, being conscious of the state might be explained in terms of the ability to avow it.

Now I will add another consideration as a premise:

5. A categorial property or event cannot be explained only by a disposition.

To see this, consider the following. To explain the glass being broken we have to refer not only to its fragility, but also to the fact that it was hit by something. Conversely, dispositions *can* (perhaps not always, but certainly sometimes) be explained only by categorial properties or occurrent states. Thus, the fragility of glass is explained by the

fact that it has a specific structure, and having such-and-such structure is a categorial property.

Now a stronger result with respect to the consciousness of occurrent mental states can be established:

6. A property of a state's or an act's being conscious cannot be explained only by any disposition, including an ability to avow or otherwise express it.

By an "explanation" I here mean the conditional of the form "if A then B," where A is a conjunction of conditions sufficient for B obtaining. Thus, while the structure of glass entails its fragility, the event of glass breaking is entailed by the conjunction of the fact that glass is fragile and the event of its being hit. The conclusion of the argument, then, can be expressed in the following way:

6*. A property of a state's or an act's being conscious is not entailed by any conjunction, whose conjuncts are only dispositions.

But Finkelstein's conditional has just this logical form. It states that the property of someone's mental state's being conscious is entailed by his ability to avow this state, and this ability is a disposition. Thus, if my argument is sound, then the neo-expressivist account of the property of being conscious of mental states and acts cannot be correct. And if neo-expressivists lack a proper account of what constitutes consciousness of mental states, then their account of self-knowledge cannot be the whole story and is dependent upon the proper account of what constitutes consciousness of mental states.

Now, a proponent of neo-expressivism might question premise 2. Thus, Finkelstein distinguishes between being conscious *that* one is in some state of mind and

being consciously in that state of mind.³² Being conscious that one fears something, say, is just being aware that one has fear, and this awareness can be obtained by, for example, a consultation with a therapist. But being consciously in fear is, according to Finkelstein, tied to the ability to express it in a self-ascription, and this ability cannot be obtained just by such indirect means. One can then argue that being consciously in fear is actually a disposition and not an occurrent property or event.³³ If this is so, my argument is unsound.

This account looks plausible when it is applied to attitudes such as fear or desire. But it immediately loses its plausibility when one tries to apply it to occurrent states or acts. Consider, first, the case of sensations, such as pain. It does seem that being conscious of one's pain involves something other than being able to avow it, namely having the phenomenal quality of pain. I might be unable to avow or otherwise express it, but it seems that if I am in pain at all, I am also conscious of it. Again, the example of paralyzed and injured person developed above can serve as an illustration of such a situation. Although it does not seem counterintuitive that a person in that situation is not conscious of his beliefs, if he has any at all, it does seem counterintuitive that a person in pain is not conscious of his being in pain. Similar considerations can be applied to other passive occurrent mental states.

In the case of mental acts the inability to express them is quite common, and one does not need to invent unusual counterexamples to illustrate it. Consider imaginings. I

³² Finkelstein 2003, 116.

³³ I am not saying that Finkelstein himself would subscribe to this argument, but this is something one can present as a defense of his position.

can lack words for the images I am now “producing” if, for example, I am trying to remember some object I saw but do not know how it is called. Yet I am surely conscious of such images. It is the same with some thoughts, as when I am trying to formulate something and still have not found the right expression to capture it. I can express a thought only after some efforts to formulate it, and these efforts are perfectly conscious. The same goes for intentional actions. If I am deeply involved in what I am doing, and especially if the action is complex and not habitual for me, I might be unable to say what exactly I am doing now until I spend some time in reflection. But of course I am conscious of what I am doing, and I am the more conscious the less simple and habitual the action is. Thus, as we see, in some cases consciousness and the ability to express are not the same, but actually impede each other.

Thus, the ability to express mental states intentionally and, in particular, the ability to express them in self-ascriptions, is distinct from consciously knowing one’s own mental states. It is striking that the recent account of self-expression developed by Mitchell Green (which was to some extent used by Bar-On in her neo-expressivist account of self-knowledge) recognizes this: Green argues that only those states can be self-expressed that can be known by introspection.³⁴ Whether they indeed can be known by *introspection* and not in some other way is, of course, the question under discussion,³⁵ but the claim that mental states can be self-expressed only if they can be known consciously is, I think, vindicated.

³⁴ Green 2007, 38-9.

³⁵ Green is not engaged in the debates on self-knowledge, so it is not a problem for him.

Note, however, that my argument leaves open the question of self-knowledge of attitudes. For all I have said, the neo-expressivist account may still be onto something in this particular case.

Criticism of the Rational Agency Approach

The first problem the rational agency approach in the Moran's version is that it can be applied only to a limited range of mental phenomena. Moran himself acknowledges that it is not suited for the explanation of passive occurrent mental states. However, even for the intentional attitudes for which the approach was designed, it is applicable only to some of them. Moran is aware that most of our attitudes are not formed by explicit rational deliberation. He argues that it is important that we have at least implicit reasons for holding them, so that a person, if asked to justify her attitude, can answer and provide these reasons. But it seems that we also have a special privilege with respect to attitudes for which we do not have rational justification.³⁶ Sometimes we hold beliefs that are not justified by the evidence we have. An extreme case is obsessive attitudes of some pathological nature. A less unusual case includes various examples of unjustified prejudices, for example racist or nationalistic ones. All such attitudes not only are not typically formed via the process of explicit rational deliberation, but also are not justified by any implicit reasons that can be adduced when the subject is asked why he holds them. This is so even for beliefs and intentions, which are the most convenient cases for

³⁶ Cf. Finkelstein 2003, 162 ff.

Moran. As for fears, desires, and so on, it is very easy to adduce multiple examples of irrational attitudes of these kinds.

Another and the main problem of Moran's approach is that it presupposes another kind of first-person knowledge, distinct both from self-knowledge with respect to our attitudes and self-knowledge with respect to our occurrent mental states, such as sensations and feelings. Namely, it presupposes knowledge of thoughts which we think during deliberation. It can be that we have, say, a rationally formed belief because in the deliberation we positively answered the question "is such-and-such the case?" But in order for us not only to *have* this belief, but also to *know* it, we need to know the answer which is the conclusion of our reasoning, and this answer is a thought. The same goes for the premises and for the intermediate conclusions we use in the deliberation: we should have self-knowledge of them to reach the final conclusion. Thus, we need to have self-knowledge of our thoughts in order to have self-knowledge of our rationally formed attitudes.

These problems give Moran's opponents a reason to charge him with a claim that he is really "engaged in a different explanatory project from proponents of the Epistemic Approach."³⁷ It is clear that the analysis of deliberative stance and transparency condition is important for understanding our practical life, perhaps for philosophy of action, but it is not clear whether this approach really explains *self-knowledge*. I will return to these problems in Chapter III. There I will try to vindicate the rational agency approach as an explanation of self-knowledge of one particular kind, at least to a certain

³⁷ Bar-On 2004, 143.

extent. I will also attempt to show wherein Moran's mistake lies. Here I conclude that its limitations are larger than Moran acknowledges: not only this approach fails to explain self-knowledge we have with respect to our occurrent sensations and feelings, but also it cannot explain self-knowledge of our own thoughts, which is necessary for the formation of our attitudes, self-knowledge of which Moran purports to explain. As I will argue shortly, this is a consequence of Moran's failure to account (and even to set a problem of accounting) for self-knowledge of our mental actions.

CHAPTER III

DIFFERENT KINDS OF SELF-KNOWLEDGE

Classification of the Objects of Self-Knowledge

We have seen that there are serious objections to all of the discussed accounts of self-knowledge. I contend that this fact is largely due to the tendency of authors to develop their accounts with preconceived assumptions or pretheoretical classifications of the phenomena with respect to which we have distinctive first-person knowledge. Thus, many of them start with the assumption that all these phenomena should be treated in the same way. This assumption is what Boyle called the Uniformity Assumption.³⁸ But there is a large variety of mental states, acts, events and attitudes, and it is far from obvious that first-person knowledge of all of them should be given essentially the same explanation. Are beliefs, sensations, feelings, moods, and thoughts really all alike? At least one should give an argument for the uniform approach to them. It is not a compelling argument that they all can be expressed in avowals, as Bar-On implies (even if they all really can be so expressed, for, as I already argued above, it does not seem natural to say, for example, that one can avow products of one's imagination, if she lacks words for the images she produces). As Rödl observes, "that avowals share certain features requiring explanation does not entail that the explanation of these features is one for all avowals. It does not entail that the concept of avowal is suited for deployment in a

³⁸ Boyle 2009, 141: "We could call this the Uniformity Assumption, for it amounts to the demand that a satisfactory account of our self-knowledge should be fundamentally uniform, explaining all cases of 'first-person authority' in the same basic way."

philosophical account of anything.”³⁹ Much of the criticism adduced in Chapter II also suggests that those approaches, which seem to be at least reasonably successful in explaining knowledge of some kinds of mental phenomena, fail as soon as one attempts to stretch them so as to explain the whole range of these phenomena.

Some authors, such as Moran, Boyle and Rödl (the last two – with an explicit reference to Kant) argue that the Uniformity Assumption is unjustified. As we remember, Moran suggested that different stories should be told about self-knowledge of passive occurrent mental states and self-knowledge of attitudes. Boyle and Rödl write about passive self-knowledge and active self-knowledge. However, while the views of these authors is much more satisfactory than the view based on the Uniformity Assumption, still they do not start with an attempt to observe the whole range of the relevant phenomena in order to classify their variety appropriately. These authors made a large step forward in rejecting the Uniformity Assumption, but still their subdivision of the kinds of self-knowledge was done without a full-fledged explicit discussion, and to start with some preconceived classification of phenomena is to leave out the possibility of errors. Of course, it is not necessary and is hardly possible to make all the relevant observations and classifications singlehandedly and from scratch, and what was already done by the authors writing on self-knowledge provides a good starting point.

Fortunately, there are works where a considerable amount of classificatory work with respect to everything mental has been already done. Surprisingly enough, a lot was

³⁹ Rödl 2007, 12. As I noticed in the relevant section, I think the concept of avowal *is* suited for a philosophical account of the communicative aspect of self-knowledge. However, it still does not help in answering the fundamental question.

done in a work which disputed the very idea of the mental, namely, in Ryle's *Concept of Mind*. Ryle has done a good job of classifying and discussing certain aspects of our sentient and sapient life that can be with more or less good reason described as dispositions to do something. Indeed, Ryle manifested a remarkable ingenuity in this endeavor. But a lot can also be learned from what Ryle has *not* done, or from what is manifestly unsatisfactory or incomplete in his work. He has more or less succeeded in subsuming motives, inclinations, agitations, moods, many of intellectual capacities under the concept of disposition. His work might be considered as preparatory in that it allows us to separate the entities with respect to which we do not have distinctive self-knowledge in a non-derivative way. As I will argue in the next section, this consideration allows us to modify Moran's position so as to reply to the criticisms raised against him. However, an attentive reading of Ryle's work also makes manifest that there are mental states that he failed to explain via the concept of disposition. These include at least sensations,⁴⁰ occurrent feelings, imaginings,⁴¹ thoughts⁴² and intentional actions (insofar as such actions involve knowledge of what one is doing). Moreover, it seems that at least some moods (such as sadness) and agitations (such as anger) do have a specific

⁴⁰ In Chapter VII of his book Ryle gives a criticism of the concept of sensation, but even if this criticism is perfectly sound, what he really succeeds in is a refutation of certain epistemological doctrines which claim that we know external objects via the intermediary of something like "sense data" or "sensations" in the sense close to "sense data". However, as Ryle himself notices (2002, 201) one may legitimately talk of such sensations as fatigue or pain. And it seems that we do have distinctive first-person knowledge of our pain.

⁴¹ In Chapter VIII Ryle criticizes the idea of imaginings as "special pictures." This, however, does not entail that there are no special mental processes we call "imaginings" and which we know with first-person authority. Sartre in (1948) and (1962) also provides a devastating critique of associating imagination with sensation, but he also builds a positive theory of imagination.

⁴² Ryle argues (Chapter IX) that we should not ascribe to our real thinking processes the names of their products, such as "judgments," "inferences" and so on. This might be reasonable, but again, it does not show that there is no thinking.

phenomenology in addition to being able to manifest themselves in either specific actions or occurrent thoughts or feelings. These might be the intermediate cases between dispositional attitudes and occurrent states and mental actions.

Now, we can reflect more on the occurrent states and actions listed above. Can we draw further distinctions between them that are relevant for our problem? I think we can. One such distinction is a distinction between passive occurrent states, such as feelings and sensations (and, perhaps, moods and agitations insofar as they include specific phenomenology), and mental actions (I will discuss what exactly counts as a mental actions in Chapter IV).⁴³ Passive occurrent states are something that happen to us rather than something we *do*. With respect to them it does not seem at least *prima facie* inappropriate that we know them by observation, although perhaps not in quite the same way as we know external objects. By contrast, thoughts, imaginings, and intentional actions are something that we in some sense do, produce, or bring about, and it seems wrong to say that we know them by observation no matter how intimate or special the observational mechanism is. I will argue for this claim in Chapter IV.

Thus, we have distinguished three main categories: passive occurrent states, mental actions, and attitudes. These categories are all distinct in the aspect which interests us, and different (although it does not mean unrelated) explanations should be given to each. I will now consider attitudes and will argue that both Moran's and neo-expressivists' approaches are helpful in explaining self-knowledge of them.

⁴³ Of course, this distinction roughly coincides with the distinction made by Boyle and Rödl, but, unlike them, I explicitly delineated dispositions. They need a separate account which I will provide in the next section.

Self-Knowledge of Our Attitudes

As I will argue here, we have self-knowledge with respect to our attitudes only in a derivative way, while the basic kinds of self-knowledge are knowledge of passive occurrent mental states and knowledge of mental acts. The understanding of this fact enables us to solve at least some problems of Moran's view.

As we saw in Chapter II, Moran claimed that even if we do not actually form beliefs or other attitudes by deliberation, it is important that we can at least give reasons why we hold it. Now, as I have shown, the wide-spread criticism of his view is that there are attitudes which we cannot justify, and yet of which we still have privileged first-person knowledge. For example, there are attitudes which are not sensitive to one's reasoning at all. Moran argues that such attitudes are in some important sense "alienated" from us, but this point does not really answer the criticism. I suggest that we look at such attitudes without preconceived opinions and see how we really get knowledge of them. Consider some unjustified belief, a belief that p . What makes it "alienated" is that even though I do not have arguments for it and, perhaps, I even know arguments against it, I still cannot help believing it. It resists my ability to change it via rational deliberation. An example of such a belief is a father's belief that his son will return from the war even if that war ended some years ago, and all the evidence points to the conclusion that he will not return. No matter what reasons are considered by the father, he persists in his belief that his son will return. In this sense it is recalcitrant to reasons. However, there are no apparent reasons to deny that he has self-knowledge with respect to this belief.

Thus, while Moran's account works with respect to rationally formed beliefs, it does not explain self-knowledge of "recalcitrant" beliefs and other kinds of attitudes. What I suggest is to use as a primary model for explanation not a rationally formed belief, but an attitude that is not formed by deliberation, perhaps even one which is insensitive to reasoning. If I will be able to provide an account of self-knowledge of such attitudes, it will be relatively easy to explain the role of rational deliberation when there is such rational deliberation.

When do I realize that I have an attitude in the first place? I think the answer is: when someone (perhaps, myself) asks me whether I believe that *p* or when I encounter another situation which forces me to use this belief consciously, that is to say, to express it either verbally or in an intentional action. Rationally formed beliefs are the result of conscious and active deliberation. Some beliefs which we do not form via explicit deliberation are still not alienated, because the process of their formation was at least partly within our view, as it were. For instance, we could check arguments for and against them at least from time to time, at least to some extent. By contrast, the formation of other kinds of attitudes is done "behind the scene," without a possibility of our conscious intervention. Because of that we do not even know that we hold this or that particular belief until we encounter a situation in which we are forced to make a judgment or an action which reveals whether we have this belief.

This conclusion is even more evident in the case of other attitudes. Consider some irrational phobia, for example the fear of spiders. How can I possibly get knowledge that I fear spiders until I encounter one? Again, I can get this knowledge only

when I encounter a situation that will show me that I have this fear. Similarly, an irrational desire will show itself to me when I either perceive the desired object or think about it. A perception or a thought will give me occasion to come to know that I have a desire for an object in question. Otherwise it is not clear how I can know that I have this desire.

One can say that attitudes give rise to behavior associated with them, and that in order to understand that we have, for example, fears or desires, we have to reflect on our behavior. The stronger the attitude, the stronger motive it will create for our behavior and the easier it will be to notice its influence and to get knowledge of it. But, of course, this kind of knowledge is not specifically first-personal and anyone other than the person who has the attitude can make the same observations and inferences as that person. This point might suggest, in line with Moran, that our “alienated” attitudes are indeed on par with external objects with regard to first-person knowledge.

However, I think that many attitudes, especially such as fears or desires, not only incline us to certain behavior, but also have specific phenomenology associated with them. We all know specific cramps of fear or urges of desire, and these cramps or urges are indeed something only we can know about ourselves. Thus, after all there is privileged first-person knowledge associated with our attitudes, but it is not knowledge of our attitudes as such, but only of phenomenology associated with them. Attitudes themselves are dispositions rather than objects that can somehow be perceived. As I already noticed, it is plausible that there are also intermediate cases between dispositional attitudes and occurrent mental states, for instance agitations and moods.

This fact, however, does not threaten the main line of my reasoning: the explanation of first-person knowledge with respect to such cases will simply include features of the explanation of passive occurrent states.

So far I have considered irrational attitudes. But what about those attitudes which are actively formed by deliberation or at least more or less “under our control”? I think that the crucial conclusion still applies to them with one important addition. These attitudes also manifest themselves in thoughts or in the behavior consistent with them in situations which show us that we have them. But these attitudes are also the result of our deliberation, that is, of thinking. Thus, we have a distinctive first-person relation with respect to these attitudes in a way derivative from our first-person relation to our thinking with which we formed these attitudes. Thus, we have first-person knowledge with respect to rationally formed attitudes in the following ways: 1) insofar as we can express them in our thoughts and intentional behavior in accordance with them; 2) insofar as we form them by explicit deliberation; 3) insofar as these attitudes have characteristic phenomenology associated with them. Of course, we can sometimes think that we believe in something without really believing it, and we can also think that we formed a belief or an intention rationally, while it was really acquired by us for some unjustified reasons, and we only later provided a rationalization for it. For these reasons self-knowledge of our attitudes is not as reliable as self-knowledge of our mental acts and passive occurrent states. Nevertheless, such self-knowledge is still distinct from knowledge we have of external objects and states of affairs.

CHAPTER IV

SELF-KNOWLEDGE OF OUR OWN ACTIONS

The considerations developed above suggest that there are two basic non-derivative kinds of self-knowledge: knowledge of passive occurrent states, such as sensation and feelings, and knowledge of our own actions. In this work I will not deal with the first kind of self-knowledge in any detail; I will only say that it is at least plausible that some sort of observational or introspective account may be adequate for treating knowledge of sensations and feelings.⁴⁴ In the rest of the work I will be discussing self-knowledge of our actions. I will argue that we possess a faculty distinct from perception which provides us with knowledge of our agency. First, however, I will have a few things to say about the historical background for such a claim.

Agency and Self-Knowledge: Some History

A number of authors have suggested that we know our actions non-observationally. The most notable historical examples include Johann Gottlieb Fichte and G.E.M. Anscombe; in our own time Christopher Peacocke has recently developed the basics of epistemology of what he calls “action awareness.” In this section I will discuss the views of Fichte and Anscombe, since their treatment of the question provides, in my view, the essential elements of the account we need.

⁴⁴ For some good arguments in favor of this view one can go to the article of A. Zimmerman “Self-Knowledge: Rationalism vs. Empiricism,” which also agrees with my approach in distinguishing different kinds of self-knowledge, albeit only two kinds.

Some of the contemporary authors who work on self-knowledge have referred to Kant for his distinctions of the “inner sense” and “apperception” which, on their view, provides the framework for the development of a non-uniform account of self-knowledge.⁴⁵ Apperception is, for them, a prototype of the kind of self-knowledge we have of ourselves as agents. While I completely agree that this is the most fruitful line of thought on the matter, the problem is how to understand this apperception. Kant himself, for instance, held that we do not possess *knowledge* via apperception. In the first *Critique* he claimed that the “I think” of apperception is not an intuition but a thought (B 157). Elsewhere⁴⁶ Kant has argued that mere thinking does not amount to experience. Thus, it does not give us experience of ourselves but merely ascribes all our experience to formally the same subject without telling us anything about that subject.⁴⁷

Fichte, however, did affirm that we have self-knowledge of ourselves as spontaneous, acting subjects. On his account we know it via what he called “intellectual intuition.” He defines and characterizes it as follows:

“Intellectual intuition” is the immediate consciousness that I act and of what I do when I act. It is because of this that it is possible for me to know something because I do it. That we possess intellectual intuition is not something that can be demonstrated by means of concepts, nor can an understanding of what intellectual intuition is be produced from concepts. This is something everyone has to discover immediately within himself; otherwise, he will never become acquainted with it at all...

...I cannot take a single step, I cannot move my hand or foot, without the intellectual intuition of my self-consciousness in these actions... Only in this way am I able to distinguish my own acting (and, within this acting, my own

⁴⁵ See Boyle 2009.

⁴⁶ In the *Reflexion* 5661.

⁴⁷ Beiser 2002, 301.

self) from the encountered object of this activity. Every person who ascribes activity to himself appeals to this intuition.⁴⁸

We see that Fichte characterizes intellectual intuition as immediate, i.e., non-conceptual.⁴⁹ Further in the text he specifies that, as with sensory perception, this non-conceptual content has to be “grasped by means of concepts” in a way similar to how ordinary sensations are conceptualized. It is impossible to simply demonstrate or explain what it is since it is a *sui generis* kind of experience, and experience is to be “discovered immediately” before any meaningful reasoning about it is possible. According to Fichte it is only via intellectual intuition that we recognize something as our own activity. In the passage cited he refers only to bodily movements; however, later in the text he writes:

Moreover, I cannot discover myself to be acting unless I also construct an image or a picture of what it is I want to produce [by acting], which image I also grasp by means of a concept. For how do I know what it is I want to produce? How could I possibly know this unless I had immediately observed myself engaged in the act of constructing a concept of a goal, that is, in a type of acting?⁵⁰

Thus, quite sensibly, Fichte includes mental actions such as imagining something in the domain of what is known via intellectual intuition.

Overall, since he actually gives place to the kind of knowledge of our own actions distinct from perceptual knowledge, I think that Fichte is a better historical ally than Kant for those who want to introduce knowledge provided by our mental agency itself.

⁴⁸ Fichte 1994, 46-7.

⁴⁹ Cf. Beiser 2002, 298

⁵⁰ Fichte 1994, 47.

In her book *Intention* G.E.M. Anscombe was primarily concerned with bodily intentional actions. As well as Fichte, she has suggested that one knows one's own actions non-observationally, and she has claimed that there are two kinds of knowledge: speculative or contemplative knowledge which is true when it is in accordance with facts; and practical knowledge, for which this correspondence goes the other way around.⁵¹ Correlatively, a mistake in the first case would be one of judgment, that is to say, would be made because one got the facts independent from him wrong. In the second case, as Anscombe writes citing Theophrastus, "the mistake is in the performance," i.e. in the failure to realize one's intention.

Anscombe has also argued that to understand practical knowledge we need to understand practical reasoning,⁵² and it seems that Moran's account of self-knowledge is an attempt to work out this connection. I have argued above, however, that while there is a lot valuable in Moran's account, it is misguided if it is understood as the whole story about self-knowledge we have with respect to our activity as rational agents. Thus, this remark of Anscombe might lead to confusion.

To sum up our historical excursus, the account of the knowledge of our actions suggested by Fichte and Anscombe includes the following essential claims:

- (1) We possess a source of knowledge different in kind from perception
(intellectual intuition or practical knowledge).

⁵¹ Anscombe 2000, 56-7. There are similar remarks in Fichte as well. Cf. Fichte 1994, 44: "It is so, because I *make* it so."

⁵² Anscombe 2000, 57.

- (2) This knowledge tells us of our own agency; moreover, it allows us to recognize our actions *as our actions* in the first place.
- (3) The truth of the judgments made on the basis of this kind of knowledge is dependent not on the accurate perception of facts independent from subject but on his own *bringing about* of the facts mentioned in these judgments, which do not exist prior to the subject's activity.

I will argue that this account is by and large correct. To do this I will criticize the rival accounts of knowledge of our own actions and then will discuss the view itself, trying to make it sound less extravagant than it does, especially in Fichte. First, however, I will discuss some fruitful objections to it, which question the necessity of the claims (1) and (2) and thus attempt to undermine the whole account.

Overt Bodily Actions and Self-Awareness

Insofar as bodily actions are concerned, the account suggested above might be criticized in a number of ways. First, one might argue that while we do indeed have a kind of knowledge of our own actions which no one else in fact has, this is due simply to our proprioceptive capacities. Proprioception is surely something we have only with respect to our own body and thus first-personal, but it might be argued that proprioception is merely another kind of perception, and thus our knowledge of our bodily actions is, after all, of an observational kind, which is contingently such that it provides us with knowledge only about our own bodies. To illustrate the contingency of the distinctively first-personal character of proprioceptive knowledge we can conceive of someone else's

brain being so hooked up to my body that she will get proprioceptive knowledge of the position and movements of my limbs.⁵³

Second, another possibility to criticize the account which favors non-observational character of knowledge of our own bodily actions is opened by the so-called ecological approach to perception, developed by James Gibson (1979) and further elaborated with the problems of self-knowledge in view, for instance, by Jose Bermudez (1998) and Johannes Roessler (2003). The important idea of Gibson was that our visual perception gives us not only exterospecific information, that is, information about our environment, but also propriospecific information, i.e. information about ourselves. This information is revealed, in particular, in the boundedness of the field of vision, the constant presence of some parts of our bodies in our field of vision, visual kinesthesia and so-called affordances, that is, possibilities for perceiving subject's actions latent in the environment. These last two sources of self-knowledge should be clarified, especially given that they are important in providing knowledge of our actions. Visual kinesthesia has to do with motion perception. The idea is that during motion the relations between variant and invariant features of visual perception "make available information about the movement of the perceiver, as well as about the environment."⁵⁴ The theory of "affordances" suggests that possibilities of various bodily actions that animals can make in their environments are not inferred from the content of their perception, but are directly perceived. Thus, the ecological view rejects the view of perception on which we (and other animals) perceive physical objects in a way which is neutral with respect to

⁵³ Cf. Bar-On 2004, 100.

⁵⁴ Bermudez 1998, 110.

the range of possible actions of perceivers. Johannes Roessler uses this approach to challenge what he sees as an objectionable assumption that underlies Anscombe's view that there are two kinds of knowledge involved in intentional action, which was expounded above. Roessler argues that what drives Anscombe to introduce the non-perceptual kind of knowledge is her acceptance of the unsophisticated view of perception which the ecological view rejects. On his view, the ecological view provides an account of perception which allows us to perceive our actions not merely *qua* bodily movements, whose descriptions alone are neutral with respect to any intention that governs these movements, but *qua* actions.

To illustrate his argument, Roessler uses examples such as driving the car (this one is borrowed from Gibson). While driving, I am looking at the road with a different attitude than a passenger of that same car does. I am not merely observing the road signs and junctures; rather, I am looking, for example, “to *see when to turn left*.”⁵⁵ That is to say, I am looking at the road not as a passive observer, but as an agent, “from within the act.”⁵⁶ Accordingly, I perceive myself as an agent who operates in this action.

While both of these sources of self-knowledge—proprioception and perception as it is understood by the ecological approach—are important and, no doubt, should be integrated into any comprehensive account of knowledge of our own actions, I think that it can be persuasively argued that there is a yet further source of action awareness. In the case of intentional overt bodily actions this source has to do with *intentional* component of the action. If, as Sebastian Rödl puts it, action is a movement that rests on a thought,

⁵⁵ Roessler 2003, 392.

⁵⁶ See, for example, Roessler 2003, 398.

then it is this thought that underlies an action (and which is, according to both Rödl and Anscombe, a conclusion of practical reasoning) that interests us here. However, to make the necessity of the treatment of this additional source of self-knowledge especially conspicuous, I will hereunder discuss specifically mental actions, in which neither proprioception nor visual perception are involved at all.

Mental Actions: Some Preliminaries

Intuitively it seems that we are active with respect to a certain part of our mental life. It seems that there are some things that we do rather than just happen to us. However, what exactly is the scope of mental agency is not so clear. Thus, before considering whether we do indeed have a special source of knowledge—that which springs from our agency itself—we need to specify what exactly will count as a mental action.

Christopher Peacocke, one of the contemporary authors who advocate the idea of non-perceptual action awareness, claims that, among other things, judgments and formations of intentions are actions.⁵⁷ Recently Galen Strawson has argued that the scope of mental action is, actually, quite restrictive.⁵⁸ He argued that it comes down to the performance of “prefatory” and “catalytic” functions. Thus, on his account only such acts as setting one’s mind at some problem, refreshing images, rehearsing formulas and the like would really be mental actions (231-2), whereas such things as judgments or deciding to do something are not. In short, “there is very little action in mental life” (244).

⁵⁷ See Peacocke 2009, *passim*.

⁵⁸ Strawson 2003.

Strawson has a point here. He is correct in that such things as formations of beliefs, at least in the vast majority of cases, are not mental actions, although the observation that beliefs cannot be formed at will is by no means new.⁵⁹ Strawson also seems to be right that remembering something is not a mental action: at least one cannot intend to remember a specific thing since in order to intend that one would have already to remember the thing in question. Perhaps, many cases of deciding something, i.e., merely acquiring a certain intention without an explicit exercise of choice, also happen to us rather than are done by us.

However, such authors as Andrei Buckareff (2005) and Alfred Mele (2009) have contended that Strawson has made an exaggeration. Buckareff claimed that Strawson failed to provide any criterion by which mental actions can be distinguished from nonactional mental events (84), and that “a good deal more mental activity than Strawson admits is actional” (83). As an example he argues that doing mental arithmetic is an action (85), with which Mele agrees. Indeed, for example, to multiply two- or three-digit numbers most people will need to do a number of actions, such as multiplying digits of one number by the other number in turn, then adding them, perhaps also rounding numbers in the beginning to facilitate calculations and also holding the intermediate results in short-term memory.⁶⁰ All these manipulations require attention, concentration, guidance, and can be done at will. And they are not merely catalysts or prefatory actions but the essential constituents of the calculation itself. Thus, Strawson is, indeed, exaggerating.

⁵⁹ See, for example, Williams 1973.

⁶⁰ Cf. Mele 2009, 32.

Unlike Strawson, Buckareff and Mele do provide a criterion by which we can distinguish mental actions from nonactional mental events. Simply put, a (mental) action is something one can try to do, where trying is making an effort to do that action.⁶¹ In other words, mental action is something one can do just because one has reasons to do it.⁶² On this account such mental events as thinking a particular thought or remembering something are, indeed, not mental actions, since one cannot try to do them. This is so because in order to try to do something, one has to have the content of what one is going to do already in her mind. There is no problem here with overt bodily actions, for it is typical for bodily actions that one understands what she wants to do before the attempt to do it. But it is not so with thinking and remembering, since if one already has the content of a thought or a proposition to remember in her mind, she has already thought or remembered it. One probably cannot also make a judgment at will since, as Strawson and many others correctly argue, they are in normal cases governed by evidence and logical laws and not by our will. One can, however, try to bring about such nonactional mental events by mental actions proper.⁶³

Mental actions proper include Strawson's examples such as drawing attention to a problem, setting one's mind and so on, but are not limited to these. They also include such genuine actions as doing calculations in one's mind, making assumptions, and imaginings, although these latter involve the operations of non-voluntary mechanisms which provide us with quasi-sensory content itself, as Strawson points out.⁶⁴ All of these

⁶¹ Buckareff 2005, 85; Mele 2009, 18.

⁶² Or something that is "responsive to practical reasons," see Dorsch 2009.

⁶³ Mele 2009, 19.

⁶⁴ Strawson 2003, 239-42.

one can do at will without mediation by other actions that will lead to their obtaining. One qualification is in order here. It seems that such actions as imagining or making assumptions do not admit the possibility of failure, and thus the very concept of trying is also problematic with respect to such actions.⁶⁵ Then one can argue that they should not even be called actions on Buckareff's and Mele's criterion. Still, making assumptions and some imaginings clearly seem to be something that we do rather than something that just happens to us. It is possible to respond to this problem in at least two ways. One can either substitute willing for trying in the criterion of whether something is an action provided above. Then a (mental) action is something one can will to do, where willing, unlike trying, does not necessarily admit the possibility of failure.⁶⁶ It is crucial that willing (and, which is more clear, trying) are not themselves actions (or else an infinite regress is threatening) but, rather, constitutive aspects of intentional actions. Those who object to the notion of willing can insist that there is a conceptual possibility of failure in such actions as imagining or making an assumption, but we are so skillful in these actions that in practice it does not really arise.

I think that such non-momentary actions as listening and watching can also be regarded as mental actions, since in these cases one needs to sustain constantly one's attention on the source of auditory or visual information respectively and sustaining one's attention seems to be an exercise of will, something that one can intend, try to do, that is responsive to practical reasons.⁶⁷ For the illustration of this claim the reader can

⁶⁵ O'Shaughnessy 2009 also considers actions for which the concept of trying seems to be inapplicable.

⁶⁶ O'Shaughnessy goes in this direction, see 2009, 171-2.

⁶⁷ Cf. Crowther 2009.

think about listening to someone who is speaking in a foreign language to which the listener is not quite accustomed or to keeping track of the conversation in a place where loud music is playing.

Such processes as mental reasoning might be considered as complex actions. Although the content of thoughts which come to mind during reasoning as well as many of the transitions between thoughts, such as logical inferences, is not something determined voluntarily, reasoning as a whole is something one can try to do and succeed (or fail⁶⁸). Such complex actions are similar to overt actions, which also involve much of what is not under an agent's control.⁶⁹ For instance, any overt action involves certain physiological processes in the organism. Yet if we are willing to call a voluntary bodily movement an action, reasoning initiated and sustained by an agent should be considered as an action too. Similar considerations should be applied to imaginings: while Strawson seems to be right that particular quasi-sensory content that arises after we make an attempt to imagine some object is not something we bring about voluntarily, this alone does not suffice to deny imagining the status of a mental action. There is plenty of genuine agential activity involved in the act of imagining: setting one's attention to do it, sustaining it, and sometimes, when the quasi-sensory content appears, manipulating it in various ways.

The question whether judgments are mental actions, as Peacocke suggests, or whether they are just passive acquisitions of beliefs, as Strawson, Mele and Dorsch

⁶⁸ What exactly a failure to reason is is not a simple matter. One sense of failure to reason is failure to reason *correctly*. Another sense, which Buckareff points out (88), is failure to reason about the specific content one intended to reason about.

⁶⁹ Buckareff 2005, 86.

(2009) think, is a difficult one. The standard argument for the claim that judgments are not actions is that they cannot be done at will since genuine judgments are supposed to be governed by evidence or, in general, by truth-related norms, and not by our will. If this is correct, then our agency with respect to judgments extends only to the preliminary and facilitatory conditions, such as keeping one's attention on the subject matter of thinking. However, I think there is one more sense in which judgments (and, for that matter, instances of reasoning such as inferences) *are* actions. Often we acquire new beliefs not merely by observing something and letting the evidence work on us (with all the facilitatory activity included) but by doing reasoning in either overt or private speech. Moreover, these utterances we do in public or *sotto voce* are, as it were, vehicles of our thought.⁷⁰ It is not always the case that we would have made the same judgments were we not uttering those phrases-vehicles. But clearly the utterances are as such actions, and those made *sotto voce* are mental actions. Thus, at least judgments made, as it were, in verbal garments, perhaps could be considered to be mental actions. However, judgments considered as acquisitions of beliefs are not actions. Neither are thoughts in the sense of entertaining new thought content, as Strawson correctly argues.

Criticism of the Rival Accounts

In this section I will consider a number of possible accounts of knowledge of our own mental actions. Some of these coincide with the views already criticized in Chapter II, but since there I have shown only that these accounts cannot function as explanations of

⁷⁰ For this idea see Soteriou 2009, 242.

self-knowledge of the whole range of the entities with respect to which we have self-knowledge, it is still to be shown that they are inadequate for the explanation of self-knowledge strictly of mental actions. I will consider (a) the observational or perceptual model of self-knowledge of mental actions; (b) the account which grounds self-knowledge of mental actions in inference of mental states, and (c) the prompting model. I will argue that all these accounts are inadequate. This result will open a way for the acceptance of the view that self-knowledge of mental actions is obtained via the source of knowledge that is parallel to and independent from sense perception, which grounds the rationality of this kind of knowledge. In the section 5 I will consider this source more closely.⁷¹

The Observational Account

The observational account of our knowledge of our mental actions will have it, as in the case with self-knowledge in general, that there is some very reliable perceptual “faculty” or “mechanism” that has our mental actions as its objects. One problem in the case of mental actions is the same that I indicated in Chapter II as a problem of the epistemic approach generally. Knowledge obtained via observation is always based on a prior identification of the object of observation. That is, if I predicate some property, say, color, of some object, say, a table, I need first to identify the object as a table. If I predicate something of myself, and if the judgment in which I make this predication is based on observation, I need to identify what I observe as related to myself. For

⁷¹ For both this strategy and for many of the arguments of this section I am heavily indebted to Dorsch 2009 and O’Brien 2003 articles.

example, I can attribute some physical characteristics to myself on the basis of observation. However, in the case of mental action we do not actually first identify something (whatever that may be) as our self, and then attribute a mental action to it.

Because self-ascriptions of mental actions are not based on identification, they enjoy what Shoemaker called immunity to error through misidentification. This immunity means that if my self-ascription is mistaken, it cannot be such because it is not I but someone else is doing the action in question. By contrast, in other cases I can see someone performing an action, make a mistaken judgment that X is performing that action, where mistake is due to the wrong identification of an agent as X. Again, one can answer this line of criticism by resorting to some version of the Cartesian view and arguing that only one self can be observed by each of us, and thus no identification is necessary. In this case such an answer is hard to refute conclusively, but there are at least two problems with it, which weaken it. First, it saddles its proponents with the dualistic view, to which most of the contemporary philosophers will raise objections on the grounds other than the problem of self-knowledge. Second, it precludes a uniform explanation of self-knowledge of mental actions and self-knowledge of bodily actions. For, in the case of bodily actions it is not only one self that can be observed as performing an action. Admittedly, this whole line of criticism is less than decisive, and yet I think it undermines the appeal of the observational account of knowledge of mental actions.

Another problem of the observational account is that it seems that when we make an assumption, imagine something, or try to listen to something while sustaining our

attention, the object of our attention is not the mental action which we do but the object of that mental action. But it seems that any kind of perception draws on the resources of attention. However, mental actions draw attention to themselves. Thus, if we had a perceptual faculty with which we observe our mental actions, it would compete with our actions themselves for the resources of attention. And so we would be the less effective in our mental actions the more aware we are about them. However, it seems that we do know about an action in which we are involved apart from knowledge obtained by second-order reflection upon our actions. And this former kind of knowledge does not seem to compete with the effectiveness of those actions. For instance, when one is carefully listening to someone else's speech, the listener knows well *that* he is listening in a way that does not impede the effectiveness of his listening, although an attempt to *say* that would indeed impede his listening.

The Inference Account

Another way to account for knowledge of our own mental actions is to suggest that we arrive at this knowledge via inference. Suppose that one tries to imagine some scenery, and then that scenery appears "in his mind's eye." If she were introspectively aware of her intention, retained the memory of that intention, then became aware of the scenery which appeared "in her mind's eye" and, moreover, noticed that these events are temporally ordered in the right way and that her imagery matched her prior intention, she can make an inference that the relationship between the intention and the resultant

imagery is not accidental. She can further infer that it was her intention that was at least part of the cause of the appearance of the image.⁷²

This account faces a number of difficulties. First, as Dorsch notices, it puts too high requirements on the agent that she must satisfy in order to know her mental actions as actions and not as mere happenings. In particular, it suggests that only those agents who possess the concepts of temporal order, accidental/non-accidental relationships or something of this kind can know their agency. Yet it is plausible that such demands are not satisfied by many subjects who do possess awareness of their agency, such as children or (some) teenagers. And it is implausible that any inference account is able to weaken these requirements enough to solve this problem.

Second, the inference view simply does not correspond to our phenomenology of mental actions. It does not seem that we ever actually do make any inferences like those suggested in the account; it seems that our awareness of our agency is much more immediate. As Dorsch argues, perhaps we do make similar inferences in some very special circumstances, but not in the normal cases.

In addition, the particular example which I have just considered highlights an additional problem. If we did not know immediately that images produced by our imagination are our own creations, it seems that we would have to face them first as indistinguishable from perceptions and to recognize them as images only later. But, as Sartre has convincingly argued, in this case we would not have satisfactory means to make those distinctions in many cases and thus would be doomed to confuse images

⁷² Dorsch 2009, 59.

with perceptions pretty often, the situation which we do not in fact observe.⁷³ It suffices to say that in normal circumstances we do not in fact experience ourselves as considering whether something is a perception or an image, and thus the question of criteria for distinguishing them does not even arise.

The Prompting Model

The prompting model suggests that judgments about the nature of our mental episodes that establish whether they were brought about by us or not are “prompted by simply paying attention to the issue, or asking oneself how a certain present mental episode has been formed.”⁷⁴ This view seems close to my own account of self-knowledge of our attitudes developed in Chapter III above. However, in the case of attitudes my account actually combined prompting with inference: on my account occasions in which one’s attitudes are exercised provide us rational basis for inference that we have those attitudes. If the same strategy is used in the case of our mental actions, then we are back to the inference account. But I have already argued that in the case of mental actions as opposed to the case of attitudes (whose knowledge is, on my view, not as immediate) this account does not work. But the proponent of the prompting model does not necessarily have to go that way. He might suggest that the higher-order judgments about our mental actions are prompted by some merely reliable causal mechanism which operates on the sub-personal level, while there are no rational relations between our mental actions and our judgments about them. This is, of course, a restricted version of

⁷³ Sartre 1948, 95-101.

⁷⁴ Dorsch 2009, 59.

the “no reasons model,” that is, the view that some class of judgments (in the present case judgments about our mental actions) is not rationally grounded at all.

As in the case with self-knowledge in general, the “no reasons model” has a problem common to any reliabilist account that our higher-order judgments about our mental actions will be rationally ungrounded and hence appear to be irrationally held. If one is opposed to reliabilist approach in general, one would need to make a compelling argument why this approach can be used only in the particular case of knowledge of our mental actions and not in other cases.

Let us suppose, however, either that the proponent of the “no reasons model” is not opposed to reliabilism in general or that he is able to provide such an argument. Still, this account faces the same problem in the case of mental actions as in the general case (discussed in Chapter II): it does seem that mental action have certain phenomenal quality (“actish phenomenal quality,” as Carl Ginet called it), that is, there is something it is “like” to be engaged in action rather than to experience something passively. For example, it seems that there is an experiential difference between actively drawing one’s attention to some sensation in the body and an involuntary switch of attention to the same sensation. Furthermore, it seems that this phenomenal quality can serve as a reason for self-ascribing an action to oneself.

Can one then claim that our higher-order judgments about mental actions are constitutive of our mental actions themselves? This claim is problematic, for it seems that children acquire the ability to self-ascribe their mental actions later than they acquire the ability to perform them. Perhaps another case when one is doing mental actions and

is not aware that one is can be encountered in some forms of schizophrenia. Some of the patients who suffer from it report the phenomenon of the so-called thought insertion, when they do not recognize their thoughts as their own.⁷⁵

Thus, I have argued that we do not know our own mental actions via observation or inference; and the “no reasons view” does not seem to be right either. Of course, no list of the possible accounts can be declared to be exhaustive on a priori grounds. However, since I do not know of other plausible rival accounts, I will assume that I have made a good argument for the view that our knowledge of our own mental actions is of a *sui generis* non-observational and yet rational kind.

Self-Knowledge of Mental Actions: More Details

Thus, since the rival accounts fail, it can be reasonably concluded that the best account of our knowledge of our own mental actions should rest on the acknowledgement of the existence of a source of knowledge independent from perception. This is just the claim (1) which I extracted from the accounts of Fichte and Anscombe (see the first section of this chapter). In the rest of the work I will, following Peacocke (2009), call it “action awareness.”

The claim (2), that is, that it is this action awareness that allows us to distinguish what is our actions from what merely happens to us seems to be also warranted by the preceding considerations. While it can be argued that proprioception and the “outer senses” (provided that one will accept the ecological approach to perception) allow us to

⁷⁵ Some discussion of these cases in relation to the problem of mental action can be founded in Proust 2009.

single out movements produced by us, it is not clear how else we distinguish, say, those thoughts or images that just come and go in our minds from those we deliberately produced or at least helped to obtain. The alternative accounts appeared to be untenable or at least weak.

The claim (3) has to do with mistakes which can occur in the exercise of action awareness. Recall that at this point Anscombe wrote that “the mistake is not one of judgment but of performance.” The problems of the *sui generis* kind of mistakes associated with action awareness are interesting, and I will discuss it in the rest of this section. The underlying idea that I will use to explicate these issues is that there is a connection between self-knowledge of our actions and practical knowledge in the ordinary sense of skill. The possibility of such a connection was, perhaps, suggested (if not quite explicitly) by Anscombe herself,⁷⁶ and it is currently discussed by some contemporary authors. However, I have doubts that it is quite intelligible to apply the concept of skill to all, especially the basic, mental actions. Therefore, I would like to indicate that what follows is only a tentative attempt to follow along these lines. Even if the account to be developed cannot be stretched to basic mental actions, it may indicate that there is a yet additional source of self-knowledge of ordinary bodily actions and, perhaps, some of the more complex mental actions (such as mental calculations). I will use as examples primarily the cases of bodily actions.

I think that in order to understand what a “mistake of performance” is it is worth to consider some criticism first. Indeed, are there mistakes of performance as opposed to

⁷⁶ Anscombe 2000, 88-9.

mistakes of judgments? Laurence Houlgate has suggested that what Anscombe called by this name is actually analyzable either in terms of mistakes of judgment upon which an agent unsuccessfully acts, or in terms of accidents that interfere with her performance and cause failures of her actions.⁷⁷ For instance, to use Anscombe's own examples, suppose a man has bought food in a supermarket according to the shopping list. Suppose, further, that he has made a mistake and bought margarine instead of butter. Houlgate argues that this mistake was based upon a mistake of judgment, namely upon a misreading of the relevant item in the list, or on forgetting, or on missing that item. In another example where a man turns right on the road after receiving a command to turn left, his mistake is based either on his misunderstanding of the command or on the confusion of his right hand with his left or on something similar, that is, on some mistake of judgment (we are also supposing that he has not ignored, disregarded or disobeyed the order, as Anscombe insisted). Another possibility of failure of performance is, according to Houlgate, connected with an occurrence of some accident. For instance, if a man has fired at a wild pigeon and a homing pigeon flew into his line of fire, this is an accident and not a mistake at all. The rejection of a *sui generis* mistake of performance, then, depends on a claim that "unless we thought that there was an error of judgment... we could not describe the case as *any sort* of mistake."⁷⁸

Notice that Houlgate did not provide an argument for *this* claim. Thus, his rejection of mistakes of performance is essentially question-begging, since he is relying on the premise that the only kind of mistake we ascribe to people includes mistakes of

⁷⁷ Houlgate 1966.

⁷⁸ Ibid., 260.

judgment. Still, it is a task of a practical knowledge theorist to explicate the notion of mistake of performance and to argue that it is not reducible to mistakes of judgment.

Let us first discuss what the possibilities of failure in performance in general are. Consider the case of an archer who shoots at the target. He may miss it because he misperceived its location, or because he wrongly assessed the strength of the wind. Failures like these are due to the agent's reliance on wrong judgments. Another set of cases includes misses which are due to such events as an unexpected rush of wind. Failures which are results of such occasions are not really mistakes at all. Quite often an agent cannot be held responsible for such failures since there are accidents which cannot be predicted in principle. Others can be prevented by taking reasonable precautions. However, it seems to me that there is also a third class of failures. There are failures which result from a poor exercise of skill by the agent whose skills are far from perfect. In the case of our archer he might miss just because he is not very skillful, even though he assesses the conditions quite correctly and nothing unusual happens at the moment of his shot.

Skill can also be characterized as knowledge how to do something.⁷⁹ Ryle has famously argued that knowledge how to perform various activities is not reducible to propositional knowledge. Even though Ryle's primary argument was recently challenged,⁸⁰ it still seems evident that skill does not consist merely in the knowledge of rules of some activity or in some similar propositional knowledge: it can be gained only

⁷⁹ For the classical discussion of "knowing how" as opposed to "knowing that" see Ryle's *Concept of Mind*, chapter 2.

⁸⁰ See Setiya 2008, 402-3.

by practice. And in any way when we are exercising some skill, especially the one which we have mastered to a considerable degree, we do not consider even those propositions, the knowledge of which is definitely part of what it means to possess that skill. All this suggests that failures which are due to the lack of skill should not be attributed to mistakes of judgment in the Anscombe's sense. Nor are they accidental: on the contrary, it is not surprising that unskillful agents make mistakes, and nothing extraordinary or even simply unexpected is needed to appeal to for the explanation of such failures.

It is not as easy to provide particular examples of mistakes of performance proper when the agent judges that she is doing something while she actually does not. I suggest that there are a few reasons for that. First, we use some of our skills, such as basic skills that operate in our everyday bodily or mental actions, at such a level that there are virtually no failures due to the exercise of these skills when there are no mistakes of judgment. Thus, we do not fail in the exercise of our skill of walking when there are no mistakes of judgment (such as failures to perceive an unevenness of the ground). Second, when we speak of exercises of more complex skills, we typically do not say of our actions that we successfully do them until we actually have finished them. Rather, we use the tentative expressions such as "I am trying to do X," which correspond to such tentative expressions as "it seems to me that p" in the case of observational knowledge. Third, (and this might be an explanation for the second point), when we master skills, we simultaneously acquire the abilities to assess our own performance in accordance with the norms which govern exercises of that skill. Still, it is possible to point out to some examples of mistaken judgments about one's own action based on the mistake in

performance. These occur when an agent still does not master a skill very well. For example, a beginner cook might think that he is currently making some really delicious food while in fact he is not (and his failure might be accounted for not by mistakes of judgment, such as ignorance of which ingredients should be used and how many of them or of the right time for some actions, but by his failure to mix the ingredients in the right way or by failure to actually perform all the actions in the right order and at the right time).

With this understanding of the distinct category of failures in performance in place we are now in the position to understand non-observational practical knowledge better. I suggest that it is related to “knowledge how.” In this my approach is close to that of Kieran Setiya, developed in his recent papers (2008) and (2009). However, as I will shortly show, I differ from Setiya in one crucial respect.

Setiya expounds his view in opposition to David Velleman’s account of practical knowledge.⁸¹ For Velleman, as for Setiya, intentions involve beliefs that one is going to do what one intends to do. The formation of these belief-components of intentions is justified by the fact that once they are formed we tend to act so as to realize them. More precisely, the belief that one is going to do A, that is involved in intention to do A, is justified by the fact that once that belief is formed one becomes aware of her own intention to do A (introspectively) and of her general ability to do A. Similarly one can

⁸¹ Velleman 1989.

justify her belief that she is currently doing A by reference to the same evidence: that she intends to do A, and that she has the general ability to do A.⁸²

Setiya's main problem with this theory is that it suggests that formation of intention involves a "leap of faith," since beliefs associated with intentions become justified only *after* they are formed, not before that. While Velleman simply accepts this consequence, it is not so appealing to many others, including Setiya. Basically, Setiya suggests to supplement Velleman's explanation by reference to the agent's knowledge how. Thus, "[k]nowing how to clench my fist is among the conditions in virtue of which I am epistemically justified in forming the belief that figures in my intention."⁸³ In general, if I understand Setiya's view correctly, the justification of one's belief that one is going to do A or that one is currently doing A is, on Setiya's account, that 1) one intends to do A, 2) one has the general ability to do A, and 3) one knows how to do A. The knowledge how is supposed to solve the problem of a leap of faith.

While I agree with Setiya that knowledge how is essential for understanding knowledge of our current actions (as well as future actions, but I will not discuss it here), I think that his account faces essentially the same problem as neo-expressivism. Since both the general ability to do A and knowledge how to do A are dispositions, they are insufficient to provide us with knowledge that we are doing A *at the particular moment of time when we are doing A*. Setiya seems to be aware of this problem and appease its acuteness in the following way:

⁸² Setiya 2008, 398-400.

⁸³ Setiya 2009, 135.

Part of what is involved in knowing how to ϕ , in the present case, is propositional knowledge that I can do it by doing ABC. That I know this proposition solves the mystery how I can know that I am doing ϕ if I know that I am doing ABC [the basic action via which ϕ is done]. But how do I know that I am doing ABC? We can say at least this. If I know how to take those basic means, this knowledge consists in the disposition to execute the corresponding intentions. Since I have this disposition, it is no accident that, when I intend and thus believe that I am doing ABC, I am doing so in fact. Exercising basic knowledge how ensures non-accidentally true belief.⁸⁴

However, while it is correct to say that if I intend to do ABC and have the disposition to execute this intention, it is non-accidental that I will do it, it does not follow that it is non-accidental that I execute the intention to do ABC at the particular moment t , when I have the belief that I am doing ABC. But then Setiya's conditions do not suffice to provide justification for beliefs about our own actions. Thus, Setiya's account appears to be insufficient.⁸⁵

However, it is not so difficult to make a correction now. What needs to be said is simply that it is not just knowledge how (together with other Setiya's conditions) that provides justification to relevant beliefs, but the intentional *exercise* of this knowledge how. Thus, my belief that I am currently doing A is justified by my intentional exercise of my knowledge how to do A. The exercise of knowledge how is not a disposition and thus can perfectly justify beliefs that one is doing something at a particular moment of time. Moreover, once the intentional exercise of knowledge how to do A is in place, it is even superfluous to specify the additional condition of being able to do A, since one who is already exercising her knowledge to do A evidently has the general ability to do A.

⁸⁴ Setiya 2009, 136.

⁸⁵ Velleman's account suffers from the same problem as well as from the "leap of faith problem," and thus is worse than Setiya's.

Perhaps, it would be helpful to specify in a more detail the concept of “intentional exercising of knowledge how.” On my account “intentional exercising of knowledge how to do A” is a cognitive aspect of the action A, which arguably is not reducible to the knowledge of propositions which express the rules and conditions of successfully performing A. However, since it is a *cognitive* aspect, it can be used to justify the belief that I am currently performing A.

This account seems to apply neatly to intentional bodily actions. Whether it is applicable to all mental actions is, as I have indicated at the beginning of this section, doubtful. An attempt to apply this account to such actions will run approximately as follows: 1) An agent’s belief that she is imagining something is justified by her intentional exercise of her knowledge how to imagine that thing. 2) An agent’s belief that she is concentrating her attention on something is justified by her intentional exercise of her knowledge how to concentrate attention.

Whether this attempt is defensible is a question for further consideration.

CHAPTER V

CONCLUSION

To summarize: I have argued that the accounts of self-knowledge which are prominent in the contemporary literature are inadequate for the explanation of all kinds of self-knowledge. I further argued that it is more reasonable to distinguish different categories of entities with respect to which we have distinctive first-personal privilege, and to explain this privilege for each of these categories separately. In this work I have suggested one classification of the relevant entities: occurrent passive mental events, mental actions and attitudes. Self-knowledge of attitudes is, as I have argued, derivative from self-knowledge of the two other kinds.

In the previous Chapter I have paid special attention to the account of knowledge of our own actions. I have argued that the most reasonable explanation of such knowledge is that we possess a source of non-observational knowledge of our actions. In principle one could have stopped at this point, but since recently some attempts were made to relate knowledge of our actions to practical knowledge in the sense of “knowledge how,” I have also presented some considerations on this point. It is plausible that “knowledge how” is related to self-knowledge of complex actions (both bodily and mental), but it is not quite clear whether the same can be said about the most simple mental actions, such as changing the focus of one’s attention.

A few words are in order to say concerning the relation between self-knowledge of attitudes and self-knowledge of mental actions. In Chapter III I have argued (in

agreement with the rational agency approach) that we have first-person privilege with respect to our attitudes partly insofar as we form these attitudes via deliberation. Now, arguably deliberation that results in the formation of beliefs (and, perhaps, also intentions, although this is much more controversial) is not something that is done at will, at least in ordinary cases. Even so, deliberation is inextricably connected with various mental actions. First, deliberation presupposes sustaining the subject's attention at its topic, which is a mental action. Second, even though judgments are arguably not actions as far as our assent or rejection of their content is concerned, the *utterances* of judgments (whether overt or covert) are actions. Arguably, these utterances are what bring the content of judgments into consciousness. Thus, mental actions make the content of our deliberations conscious, providing self-knowledge of our thoughts and reasoning, as well as, derivatively, of self-knowledge of our attitudes.

REFERENCES

- Alston, W. 1971. "Varieties of Privileged Access." *American Philosophical Quarterly* 8: 223-241.
- Anscombe, G. E. M. 2000. *Intention*. Cambridge, Mass: Harvard University Press.
- Armstrong, D.M. 1994. "Introspection." 109-117 in Q. Cassam ed. *Self-Knowledge*. Oxford Readings in Philosophy. Oxford: Oxford University Press.
- Armstrong, D.M. 1963. "Is Introspective Knowledge Incorrigible?" *The Philosophical Review* 72(4): 417-432.
- Bar-On, D. 2004. *Speaking My Mind: Expression and Self-Knowledge*. Oxford: Clarendon Press.
- Bar-On, D. and Long, D. C. 2001. "Avowals and First-Person Privilege." *Philosophy and Phenomenological Research* 62: 311–335.
- Beiser, F. C. 2002. *German Idealism: The Struggle against Subjectivism, 1781-1801*. Cambridge, Mass: Harvard University Press.
- Bermudez, J. L. 1998. *The Paradox of Self-Consciousness*. Cambridge, Mass: MIT.
- Boyle, M. 2009. "Two Kinds of Self-Knowledge." *Philosophy and Phenomenological Research* 78(1): 133-164.
- Buckareff, A. 2005. "How (Not) to Think about Mental Action." *Philosophical Explorations* 8(1): 83-89.
- Cassam, Q., ed. 1994. *Self-Knowledge*. Oxford Readings in Philosophy. Oxford: Oxford University Press.

- Crowther, T. 2009. "Perceptual Activity and the Will." 173-191 in L. O'Brien and M. Soteriou eds. *Mental Actions*. Oxford: Oxford University Press.
- Dorsch, F. 2009. "Judging and the Scope of Mental Agency." 38-71 in L. O'Brien and M. Soteriou eds. *Mental Actions*. Oxford: Oxford University Press.
- Fichte, J. G. 1994. *Introductions to the Wissenschaftslehre and Other Writings, 1797-1800*. Translated and edited by Daniel Breazeale. Indianapolis: Hackett Pub. Co.
- Finkelstein, D. H. 2003. *Expression and the Inner*. Cambridge, Mass: Harvard University Press.
- Gibson, J. J. 1979. *The Ecological Approach to Visual Perception*. Boston: Houghton Mifflin.
- Green, M S. 2007. *Self-Expression*. Oxford: Clarendon Press.
- Houlgate, L. 1966. "Mistake in Performance." *Mind* 75 (298): 257-261.
- Martin, M.G.F. 1998. "An Eye Directed Outward." 99-122 in B.C. Smith, C. Wright, and C. Macdonald eds. *Knowing Our Own Minds*. Mind Association Occasional Series. Oxford: Clarendon Press.
- McDowell, J. H. 1994. *Mind and World*. Cambridge, Mass: Harvard University Press.
- Mele, A. 2009. "Mental Action: A Case Study." 17-37 in L. O'Brien and M. Soteriou eds. *Mental Actions*. Oxford: Oxford University Press.
- Moran, R. 2001. *Authority and Estrangement: An Essay on Self-Knowledge*. Princeton, N.J.: Princeton University Press.

- Moran, R. 2002. "Frankfurt on Identification: Ambiguities of Activity in Mental Life." 189-217 in S. Buss ed. *Contours of Agency: Essays on Themes from Harry Frankfurt*. Bradford Book, Cambridge Mass: MIT Press.
- O'Brien, L. 2003. "On Knowing One's Own Actions." 358-382 in J. Roessler and N. Eilan eds. *Agency and Self-Awareness: Issues in Philosophy and Psychology. Consciousness and Self-Consciousness*. Oxford: Clarendon Press.
- O'Shaughnessy, B. 2009. "Trying and Acting." 163-172 in L. O'Brien and M. Soteriou eds. *Mental Actions*. Oxford: Oxford University Press.
- Peacocke, C. 1998. "Conscious Attitudes, Attention and Self-Knowledge." 63-98 in B.C. Smith, C. Wright, and C. Macdonald eds. *Knowing Our Own Minds*. Mind Association Occasional Series. Oxford: Clarendon Press.
- Peacocke, C. 2009. "Mental Action and Self-Awareness (II): Epistemology." 192-214 in L. O'Brien and M. Soteriou eds. *Mental Actions*. Oxford: Oxford University Press.
- Proust, J. 2009. "Is There a Sense of Agency for Thought?" 253-279 in L. O'Brien and M. Soteriou eds. *Mental Actions*. Oxford: Oxford University Press.
- Rödl, S. 2007. *Self-Consciousness*. Cambridge, Mass: Harvard University Press.
- Roessler, J. 2003. "Intentional Action and Self-Awareness." 383-405 in J. Roessler and N. Eilan eds. *Agency and Self-Awareness: Issues in Philosophy and Psychology. Consciousness and Self-Consciousness*. Oxford: Clarendon Press.
- Ryle, G. 2002. *The Concept of Mind*. Chicago: University of Chicago Press.
- Sartre, J.-P. 1962. *Imagination: A Psychological Critique*. Ann Arbor: University of

Michigan Press.

Sartre, J.-P. 1948. *The Psychology of Imagination*. New York: Philosophical Library.

Schooler, J.W. 2002. "Re-Representing Consciousness: Dissociations Between
Consciousness and Meta-Consciousness." *Trends in Cognitive Sciences* 6: 339–
344.

Setiya, K. 2008. "Practical Knowledge." *Ethics* 118 (3): 388-409.

Setiya, K. 2009. "Practical Knowledge Revisited." *Ethics* 120 (1): 128-137.

Shoemaker, S. 1998. "Introspection and the Self." 118-139 in Q. Cassam ed. *Self-
Knowledge*. Oxford Readings in Philosophy. Oxford: Oxford University Press.

Smith, B. C., Crispin Wright, and Cynthia Macdonald. 1998. *Knowing Our Own Minds*.
Mind Association Occasional Series. Oxford: Clarendon Press.

Soteriou, M. 2009. "Mental Agency, Conscious Thinking, and Phenomenal Character."
231-252 in L. O'Brien and M. Soteriou eds. *Mental Actions*. Oxford: Oxford
University Press.

Strawson, G. 2003. "Mental Ballistics or the Involuntariness of Spontaneity."
Proceedings of the Aristotelian Society 103(3): 227-257.

Velleman, J. D. 1989. *Practical Reflection*. Princeton, N.J.: Princeton University Press.

Williams, B. A. O. 1973. "Deciding to Believe." In *Problems of the Self; Philosophical
Papers 1956-1972*. Cambridge: Cambridge University Press.

Wright, C. 1998. "Self-Knowledge: The Wittgensteinian Legacy." 13-46 in B.C. Smith,
C. Wright, and C. Macdonald eds. *Knowing Our Own Minds*. Mind Association
Occasional Series. Oxford: Clarendon Press.

Zimmerman, A. 2008. "Self-Knowledge: Rationalism vs. Empiricism." *Philosophy Compass* 3(2): 325-52.

VITA

Name: Anton Sergeevich Kabeshkin

Address: 314 Bolton Hall
4237 TAMU
College Station, TX 77843-4237

Email Address: akabeshkin@gmail.com

Education: B.S., Information Technology and Economics, South Ural
State University, 2005
M.A. Philosophy, Texas A&M University, 2011