CHAPTER

14 On Being Someone 🔓

J. T. Ismael

Abstract

Of the things that happen, a person thinks of some as things he does. What distinguishes these from those that merely happen to him? The things a person does are produced by volition. But that just pushes the question back a step. If a person's actions are his doings because of their connection to volition, what about the volitions? This chapter suggests that the answer has to do with the logic of the choice situation. It argues that the practical judgments arrived at as the result of deliberation are non-question-begging and inalienably up to a person, that there is no way to take a passive attitude toward them, treating them in the way he treats the objects of other beliefs as something that is there to be represented.

Keywords: action, choice, deliberation, doing, judgment, volition

Subject: Philosophy of Action, Philosophy, Philosophy of Mind

Collection: Oxford Scholarship Online

WHAT IS THE difference between the things that I do, like walk across a room and pick up a newspaper, and things that my body does, like regulate the flow of blood and marshal white blood cells to the scene of injury? Or what is the difference between a wink or a blink, or diving and falling? The first of each of these pairs is something that I choose and initiate. The second is something over which I have no control. That is about as straightforward a thought as can be. But what is this I that is doing the choosing and that is thought to be the source of action. Where is it located?

If we want to appeal to the role of the self to draw the distinction between action and mere behavior, we need a literal interpretation of that talk that meshes with what our neuroscience teaches us about what is actually going on in the brain. But it has become popular among some who study the mind scientifically to dismiss the idea of the self as an illusion, an understandable but immature idea to be swept aside with the advancing understanding of the brain. Dennett, Metzinger, Gazzaniga, and others draw what they see as the natural conclusion from the neuroscientific examination of the brain, namely, that there is no self.[1] Various reasons for this claim are given. But one that recurs again and again both in print and conversation is that there is simply nothing in the head that we can assign as referent of "I". As Dennett says

> The revisionist case is that there really is no proper self: . . . ["I" fails to] correspond to anything that actually exists in one's head.[2]

The claim is that the idea of the self as the thing that controls my behavior when I act has no application at the level of neuroscience. Looking at the brain through the neuroscientist's eyes we see neurons sending signals to other ↳ neurons. We can make some topological and functional divisions, but there doesn't seem to be a fused locus of control where the threads come together and from which action issues. The neuroscientist looking in the brain for the self finds himself in the position that Dennett describes jokingly here:

> On my first trip to London many years ago I found myself looking for the nearest Underground station. I noticed a stairway in the sidewalk labeled "Subway," which in my version of English meant subway train, so I confidently descended the stairs and marched forth looking for the trains. After wandering around in various corridors, I found another flight of stairs, leading up, alas, and somewhat dubiously climbed them to find myself on the other side of the intersection from where I had started. I must have missed a turn, I thought, and walked back downstairs to try again. After what seemed to me to be an exhaustive search for hitherto overlooked turnstiles or side entrances, I emerged back on the sidewalk where I had started, feeling somewhat cheated. It finally dawned on me that a subway in London is just a way of crossing the street underground. Searching for the self can be somewhat like that. You enter the brain through the eye, march up the optic nerve, round and round in the cortex, looking behind every neuron, and then, before you know it, you emerge into daylight on the spike of a motor nerve impulse, scratching your head and wondering where the self is.[3] We might call this the "there-is-nothing-in-the-head-that-corresponds-to-the-self" problem.[4]

1. What Am I?

The problem turns out to be related to an issue emphasized by Descartes in one of his most powerful arguments for dualism. Descartes asserted that we have immediate phenomenological apprehension of the self as a thinking thing and that the self, so apprehended, is simple and indivisible. He writes,

> When I consider the mind, that is to say, myself inasmuch as I am only a thinking thing, I cannot distinguish in myself any parts, but apprehend myself to be clearly one and entire.[5]

And he argued, if the self is not made of parts, it cannot be made of matter because anything material has parts.

> There is a great difference between a mind and a body, because the body, by its very nature, is something divisible, whereas the mind is plainly ↳ indivisible....[I]nsofar as I am only a thing that thinks, I cannot distinguish any parts in me.. . . Although the whole mind seems to be united to the whole body, nevertheless, were a foot or an arm or any other bodily part amputated, I know that nothing would be taken away from the mind. (1980, 97)

The two problems are related. The first suggests that the idea that there is something in the brain-body machine that plays the role of the self as an agent of action is belied by a close look through the eyes of the neuroscientist. The second makes explicit why none of what the neuroscientist sees adds up to a self. When we look at the fine-scale material structure of the brain, we find neurons connected in complex networks that regulate the voluntary and involuntary movements of the body, but no thing—no indivisible nugget of selfhood—that receives the signals coming in from the senses and orchestrates the motor activity.

2. Self-Governance

I think that both of these complaints embody a confusion about what kind of thing the self is, considered in its role as agent. The plan in what follows is to begin by looking at other examples of systems that support the vocabulary of agency and allow us to make the distinctions that we want to make between action and mere behavior. The kind of system that I have in mind is one that I call a self-governing system. Self-governance is a type of dynamical organization. We can find actual and idealized examples of self-governance writ large in systems like companies, sports teams, and democratic countries as well, and by looking at these examples, we can get some insight into the question: What kind of thing is this object—the I—that acts when I act, doing what I do?[6] What kind of unity does it have? And is it the kind of thing that the neuroscientist should expect to see when he looks at the brain? Consider different kinds of complex system, starting with the most general category.[7]

Simple groups: simple groups are mereological sums of components. Any arbitrary collection of things, from the dust particles on a shelf, the set of things whose English names begin with R, or the set that contains the Eiffel Tower, the coins in my pocket, and Arnold Schwarzenegger's big toe, forms a simple group. There is nothing more interesting to say in general about the dynamics of a simple group than there is about the activity of the parts individually. There is not in general a more compact description of the dynamics of the whole than what is obtained by describing the activity of the parts and conjoining them.

p. 277 **Dynamical systems:** dynamical systems are systems of components that are worth treating as a dynamical unit because there are constraints on joint behavior that make for a relatively simple dynamics. They typically consist of a collection of parts bound by mechanical, chemical, or electromagnetic bonds so that there is restriction on their relative motion resulting in a reduction in the number of degrees of freedom relative to what we would get by simple conjunction of dynamics of their constituents.[8] Consider, for example, a steam engine. The engine is made of a huge number of particles bound together in a way that constrains their relative motion. The way that we model such systems is usually by building information about the fixed structure of the system into its configuration space using a small number of parameters to specify state-dependent properties. The dynamics is then given as a function over a relatively low-dimensional configuration space. Someone learning how steam engines work doesn't need to look at the full configuration space for all of the particles that jointly comprise the engine, a space with 6n dimensions, where n is the number of particles the engine contains. She need only keep track of the relative positions of the moving parts, a space of around 12 dimensions, going by the diagram in Figure 14.1.

Figure 14.1
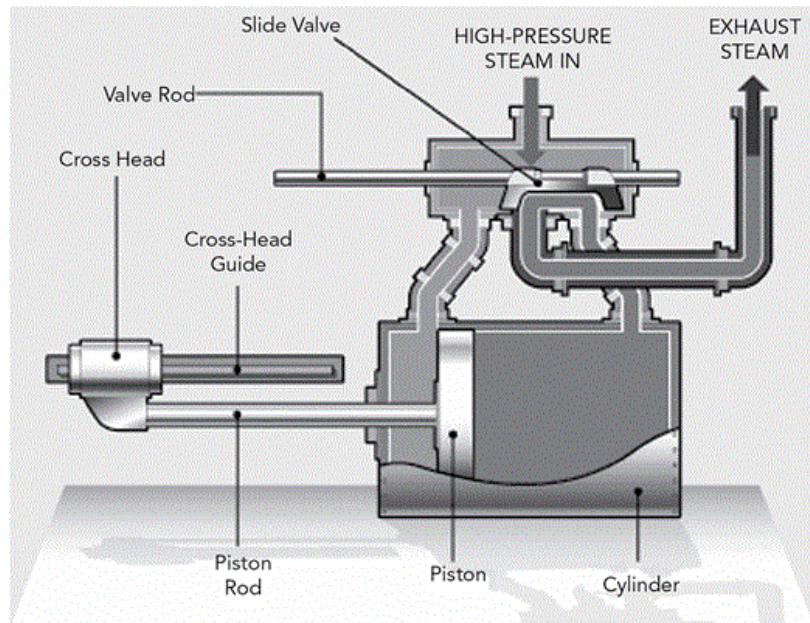
How steam engines work.

Self-organizing systems: self-organizing systems are a special class of dynamical systems. The hallmark of self-organization is the appearance of order in a system of interacting components without any centralized control. There is dispute about whether there is a general dynamical characterization that covers all instances, but central examples include termite colonies, schools of fish, unregulated crowds, traffic systems, and free market economies. In these cases there is no real centralization of information or control, but the actions of each affect the others in a manner that produces an overall appearance of deliberately coordinated activity.[9] Think of how people responding to other people's movements in a crowd can make it seem that a crowd has a mind of its own. A great deal of research has been devoted in the last 30 years or so to understanding mechanisms of self-organization.

Self-governing systems: self-governance contrasts with self-organization. In a self-organizing system, all behavior is emergent from the aggregated activity of components, each doing its own thing. The coupling among components can generate the surprising appearance of coordination, but there is not really any pooling of information and centralized control of activity. In a self-governing system, by contrast, there is some centralization of information and some top-down regulation of behavior. I will put this by saying that there is a collectivization of epistemic and practical effort among components. Information distributed throughout the systems is collected and synthesized and used as input to a decision procedure aimed at the collective good, in the precise sense that the utility function that goes into the decision-theoretic computation is one that defines the good for the system as a whole.[10] Think of the difference between a centrally controlled army and a loosely organized band of rebels operating independently, or the difference between a society without a publicly enforced government and a society regulated by institutions that are designed to give voice to the will of the people through polls and referenda or elected representatives, and implement policies that carry out that will. On this view, the government is not a separate system that imposes its will on the people. It is rather the machinery whose purpose is to express and implement the people's collective will.[11] If we were to give a formal rendering of the difference between self-organizing and self-governing system, the crucial difference for our purposes is that in a self-governing system there is both an epistemic standpoint that synthesizes the collective knowledge and a systemwide deliberative standpoint that plays some role guiding the activity of the system in which the collective good appears explicitly as a term in the utility calculation.[12] The deliberative procedure recognizes and weighs claims on

behavior put forward by subsystems (these can take any number ↳ of different forms: in human agents, the motley set of drives and appetites cobbled together by nature; in a polis, the desires of individual citizens) and makes an all-things-considered judgment about what is best for the system as a whole.

How much collectivization there is on both the practical and epistemic side, and the role that the systemwide viewpoint plays in regulating behavior, is as varied as the role that the governing board of a company, the government of a country, or the commanding structure of an army plays in regulating the behavior of the whole. Some companies are ruled by an iron hand of a governing board that micromanages every activity. Others treat their parts like semi-autonomous branches that check in periodically for gentle guidance from a company-wide plan.

In the last decade or two, with the explosion of research on complexity, self-organization has figured increasingly in discussions of human behavior. There was a time when all human behavior was thought to be coordinated by a central intelligence. We have come to appreciate that there's a great deal of human behavior that is the product of self-organization.[13] It is certainly the case, for example, that the beating of the heart, the activity of the immune system, the focusing of the eyes, and many of the processes associated with balance and locomotion are the product of self-organization. But not all human behavior is the result of self-organization. Some of it is the product of the top-down control of self-governance. The details of how self-governance is implemented in the brain are not fully understood, but no one—not even Dennett or Metzinger, when you look past the analogies with self-organizing systems, and at the details of their accounts—denies that there is the genuine forging of a collective deliberative standpoint in the human psyche that can play some role in the determination of behavior.

What the deliberative standpoint is supposed to do is take appetites and drives (prima facie claims on behavior, but which often conflict with one another and overarching plans and projects) as input, weigh them against one another and larger aims, and (in the light of all that is known about the world and the system itself) make an all-things-considered judgment about what is best. This all-things-considered judgment is then adopted as a collective plan. To say this is not to deny that much of our day-to-day lives is occupied by practiced routines that are conscious, but nondeliberative, or that there is perhaps less genuine self-government and more post-hoc rationalized self-organization than we might have thought in our lives. But I am strongly inclined to say that what makes those activities ours—that is, what makes it the case that they are things that we do, rather than things (like the digestion ↳ of food or the white blood cells) that our bodies do—is that they are available to the deliberative standpoint and can be brought under its control.[14] And I say that a self-governing system is constituted as a practical agent in its own right by the activity that collectivizes deliberation.

3. Unity

So a self-governing system is a system in which there is an internal reflective standpoint and it is the possession of that internal reflective standpoint that gives it several different kinds of unity that systems like engines or ant colonies don't possess. The first is what I call the synthetic unity attained when information drawn from incommensurate sources is mapped into a common frame of reference. The second is the univocity attained when a set of separate, potentially conflicting information streams is united into a collective voice. The third is the dynamical unity achieved when the parts of a system operate under the command of a single voice. These three types of unity are linked in interesting ways and I'll try to bring out some of the connections.

3.1. Synthetic Unity

A collection of parts pooling information to form a collective representation about a matter of common interest, for example, could be a set of detectives interviewing different witnesses and investigating different parts of a crime scene and meeting to iron out a consistent story about what went on, or the crew on a ship taking measurements from different spots and meeting in the central cabin to combine results to create a fix on the ship's location. In these cases, we have separate informational streams with partially overlapping content being combined into a single overarching representation that can then be put to use forming policies or guiding action. One doesn't combine informational streams by just having them dump into a common bin. The information has to be integrated, or rendered commensurate. A uniform descriptive vocabulary and a common scheme of singular reference have to be provided. In the example of detectives weaving a coherent narrative out of the partially overlapping testimony of witnesses, personal time lines and spatial movements are plotted jointly in an objective frame of spatio-temporal reference to identify points of contact. The time at which one witness was watching the late show is identified with the moment

another saw a stranger in the alley, a bang heard by one is connected to a flash seen by another, a glove ↳ reported missing by one witness is identified with one found later at the scene. Testimony from one witness has to be related to and reconciled with testimony from other witnesses. As a consistent narrative is constructed, information is added where narrative holes emerge; trajectories are continued when they pass out of sight; adjustments are made to resolve conflicting accounts. The integrated story doesn't report something seen by any one witness. That's not its role; it is a reconstructed compilation that identifies points of contact and reveals extensional relations between them.

The mind faces the same task in integrating information coming in through different sensory pathways. The integration is performed by subpersonal processing that takes in sensory information drawn from different modalities and represented in incommensurate ways and issues a unified conception that allows us to relate the content of different streams to one another, fill in wholes, resolve conflict, and so on. Integration holds a number of practical advantages. It allows the senses to share information: we can raise and answer questions about how the object felt relates to the one seen, where in visual space the awful smell is coming from, the direction in which to walk to reach the source of the noise, and so on. There are confirmational benefits too; the contents of different streams can reinforce or undermine one another. When we have different sources of overlapping information, we can more easily identify errors, distinguish information from noise, and be more confident in the parts that mutually support one another.

3.2. Univocity

Synthesis is a purely information-theoretic notion involving the resolution of multiple informational streams consolidating information distributed across the system into a single coherent stream. It doesn't matter for those purposes that the resulting stream takes a linguistic form. On a ship, for example, the information might be used to generate a map; the team of detectives might employ a graph or timeline to depict the unfolding of the crime. But when a complex system resolves a collection of potentially competing informational streams into a single coherent stream, and that stream has linguistic expression that allows for the explicit self-ascription of intentional states, then that system has what I will call "a voice." What I mean by this is best conveyed by illustration. Consider the way that elections and referenda turn the cacophony of competing individual voices in a population into univocal judgments on matters of public concern. A population is officially polled and the results of polls act in an official capacity as expressions of

the collective will of the ↳ people. The electoral process gives the community a collective voice. We can speak in a loose and metaphorical way of the voice of the people, but without the electoral machinery that resolves the collection of individual opinions into a collective "yes" or "no," there is no truly collective opinion.[15]

Or consider the procedures that turn the opinions of Supreme Court justices into a single collective opinion, expressed in the form of a written or spoken pronouncement issued by the collective. An even better example is provided, perhaps by juries. Without collective pronouncements, there is the collection of opinions but no opinion of the collective. When we say "The people have chosen so and so as their new president," "The court has decided that the amendment to the law is unconstitutional," or "the jury proclaims a verdict," the attitudes in question are attributed to the collective. The choices of the people are not mine or yours; they are ours. The opinions of the Court do not belong to Souter or Roberts; they are property of the group.[16] The determination of the jury is not the determination of any individual member of the jury; they belong to the whole. Collective opinion can, but need not be, obtained by applying a simple mathematical function to the opinions of its members. Indeed, in the typical case the process that generates the collective opinion—the back and forth of debate leading up to an election or a ruling—has a complex dynamics involving feedback between the opinion of the individual and the opinion of the group. Individual opinion partly constitutes the opinion of the group, but it is also affected by the way it takes the group to be leaning, so that the two are locked in a relationship of continuous reciprocal influence, making them effectively unpredictable. We see this in the swings of popular opinion leading up to an election and in the way the votes of jury members change in response to straw votes.

When it is given a collective voice, a complex system constitutes an intentional system in its own right. And that status carries with it certain responsibilities. Above all, it carries with it the responsibility to try to meet the demands of rationality, that is, to try to ensure the consistency of one's attitudes and their closure under the logical operations. We never entirely satisfy those demands. Persons, juries, and courts often issue inconsistent pronouncements, and none of us is really fully aware of all of the logical implications of our judgments, but there are systems in place to check for inconsistency and incompleteness, and when either is noticed, procedures are undertaken to correct them. More importantly, a system with a collective voice is committed to consistency and therefore can be faulted for inconsistency. We can't fault Supreme Court justices for disagreeing with one another on a personal level, but when they start self-ascribing judgments

p. 283 in a collective voice, they ↳ commit themselves to the consistency of the collective judgments. If they affirm that p, they ought not to affirm that ~p. If they affirm that p and q, they are committed to (p&q), and so on.[17]

The human mind has a collective voice in this sense. It not only integrates sensory information, it explicitly self-ascribes intentional states. Let me say a little bit about why the explicit self-ascription of intentional states is important. It has been frequently remarked that one of the things that distinguishes humans from lower animals is that we have language, and language gives us the ability to think about thinking. The idea here is that language is special in giving us a way to represent our representational states. This allows us to make them objects of attention, and this in its turn allows us to view them as a group and raise questions about their second-order properties. Are they consistent? Do they cohere? Can we provide reasons for thinking they are true? When we express our representational states in linguistic form and self- ascribe them in "I think" thoughts, we can check them for consistency and logical closure, implement procedures designed to bring them closer to a logically consistent and closed set. This makes the human mind a very special sort of self-regulating representational system. Our ability to think about thinking, to draw inferences, to check for consistency and correct errors is an expression of a commitment to meet the demands of rationality. Any system that is expressing itself linguistically and self-ascribing intentional states in "I (or we) think that p" sentences is going to share this feature with the human mind.

There is a nontrivial transition from having an internal metacognitive perspective to taking oneself as subject of those attitudes in a manner that recognizes the possibility of other subjects with attitudes different from our own. It has been argued that this transition brings with it new norms that emerge in a specifically social dialectic that involves the exchange of reasons.[18] That transition is one that is certainly made by the time one is communicating and expressing oneself linguistically. The attitudes self-ascribed in

the first person in such a setting—"I believe/desire/hope/will that p"—are made in the mind's collective voice, and they belong to the mind as a whole as surely as those self-ascribed by the spokesperson for the Supreme Court belong to the justices as a group, or the verdicts of the jury belong to the collection as a whole. In self-ascribing attitudes, the mind is affirming attitudes whose coherence as a group it also thereby commits itself to. Affirmation is a little different from reporting that such and such is the case. Reporting is a form of description; it carries the implication of a subject matter that obtains independently

p. 284 of its pronouncement. Affirmation (or "avowal") has a performative character; ↳ it makes true what it reports by pronouncing it.[19] The people haven't spoken until election results are in, the jury hasn't "made up its mind" until it has rendered a verdict, and the Supreme Court hasn't made a decision until that decision is self-attributed on behalf of the collective by its spokesperson—until the spokesperson says "we, the jury, or we the justices of the court, affirm that. . ."[20] Performances of this sort are truth bearing (what they say is true) but self-fulfilling (they make it true by saying it).[21] Making up your own mind—judging that p or willing that q—is not a report; it is an avowal.[22] When I judge or will that p, I commit myself to it on a personal level and to everything that follows from it in conjunction with other personal commitments and all of the norms that that brings in train.

The kind of unity that is possessed by subjects of intentional attitudes—the "I" of "I think that p"—is not given, but achieved, and the suggestion here is that it is achieved by forging a collective voice.[23] I think that this way of understanding things properly captures both the sense in which I—that is, the subject of these self-attributed thoughts and impressions, the thing that thinks when I think and acts when I act—am simple, and the sense in which I am complex. I am complex because I am composed of a collection of subpersonal components, but I speak to the world with a single voice. My voice doesn't decompose, and the attitudes self-attributed in my personal voice belong to me and not my parts.[24] Voices are not made of voice parts, and the attitudes self-attributed by voices aren't attributed to any part of the system that produces them. Mereology has become entrenched in philosophy as the right way of talking of the relations between parts and wholes, but it does not give the compositional logic of voices. A community is a collection of people, but a collective voice is not a collection of voices.[25] To what or whom does the voice belong? To the collective. When we have a system that not only integrates informational streams into a unified model of the world, but issues judgments in a collective voice, we have a new nondecomposable unit that can support the attribution of intentional states.

When a system develops an internal point of view and starts making judgments and decisions, the unity that emerges is a real unity, not the as-if unity attributed to the anthill by the curious spectator. The system is constituted as a subject of judgment and volition—that is, a system that engages in epistemic and practical reasoning and issues judgments and self-commands obeys epistemic and practical norms—by the collectivization of epistemic and deliberative activity. Enforcing the norms means possessing an internal metacognitive awareness, so a self-governing system has a reflective perspective. It represents its

p. 285 intentional and volitional states, checks them for consistency and ↳ coherence, spends time reflecting on its beliefs and values, and develops principles for resolving conflict. You can see the potential here for a hierarchy of self-regarding attitudes, that is, attitudes to its own intentional and volitional states.[26] How much of that hierarchy is realized will depend on the system in question. A football team or an army might engage in collective deliberation without much reflection of this kind. In human agents, the hierarchy of self-regarding attitudes is quite developed. The "I" of a normal human adult is the "I" of the reflective standpoint.

3.3. Dynamical Unity

The dynamical benefits of synthesis are easy to discern. A system that is pooling information can put the integrated body of information to use to form a coordinated plan of action. A system that has a collective voice can also check its commitments for consistency, articulate goals and reason instrumentally about how to achieve them, and engage in all kinds of self-regulating behavior. But there is more. A collective voice turned inward not only affirms but also commands; feeding back into the lower-level organization of the system, it can guide the behavior of the parts of the system. It is in this capacity that the intentional subject —the "I" of "I think"—serves as a subject of volition, the "I" of "I do." This happens in social collectives, too. Think of how laws voted in collectively by a population constrain the activity of its members. Or think of a committee that comes together to decide its collective activities and then disbands, leaving each member to carry out its part of the collective plan. As a general phenomenon, a dynamical link from higher to lower levels of organization is the source of most macroscopic order. There are channels for the propagation of information between levels in all kinds of complex systems, simple causal links that allow the global state of a system to affect its components.[27] What is really special about intentional agents is that the parts of a complex system under the command of a collective voice act with a singleness of intent and purpose that is impossible for the collection of components acting alone.

A collective voice has an external role too. Given public expression, it can mediate the communication between the parts of a system and other systems at the same level of complexity. When this happens, the systems in question form a network in which each of them is a node. A person is a node in a social network. A corporation is a node in a business network. A nation is a node in an international community. And when

p. 286 we have a system of collectives communicating through public voices we get an emergent dynamics ↳ at the intercollective level: special patterns of interaction that are relatively indifferent to the dynamics at the lower level often not predictable from the laws that govern their components. A complex system in a public setting with a voice-piece making public assertions on its behalf has the status of agent, with all of the commitments and entitlements that that status carries in the relevant kind of network.[28] This goes for governments and corporations as surely as it does for persons. Populations band together into national units with governments acting as voice-pieces, giving rise to a dynamics at the international level. Complex interactions involving feedback and feedforward can make the dynamics at the intercollective level largely autonomous of, and effectively irreducible to, the specific activities of population members. The unified voice of the internal monologue does both of these things; it allows us to operate as unified agents in interaction with other selves—loci in a social network—and it also exercises some control over its components, consolidating resources and allowing a collection of sensorimotor subsystems to make a coordinated effort in pursuit of a common goal.[29]

The overall suggestion, then, is that forging a unified standpoint that collects and organizes information from a number of sources gives rise to a subject whose point of view spans those of the sources, and forging a unified voice in which judgments are affirmed and decisions pronounced gives rise to an intentional subject to whom those judgments and decisions belong. The minds of many lower animals consist of a collection of quasi-independent subsystems each with its own sensorimotor loop that responds selectively to a particular class of stimuli and regulates a specific set of behaviors. When these kinds of sensorimotor loops are cobbled together in the right way, the system as a whole can exhibit highly adaptive behavior, but if there is no integration of sensory information, there is no point of view that spans the sensory modalities, no "I" to which visual, auditory, tactual, and other states can all be ascribed. And if there is no explicit self-ascription of intentional states and attempt to make them as a group into a logically disciplined set, there is no "I" to which those states are properly jointly ascribed. What distinguishes a system like this from the human mind is, among other things, the integration of sensory information. The existence of a perceptual subject that occupies a point of view that spans the modalities, or spaces, is an artifact of the integration process.

In a self-governing system, there is some one making the choice, some one exercising control, some one making judgments and undertaking commitments. The someone here is not an individual substance or material particular lurking inside the system, it is a point of view (or, if you like, is the formal subject of a point of view, the occupant, or the possessor), and it has a kind of ↳ formal one-ness that is the product of, rather than a precondition for, the collectivization of epistemic and practical deliberation.[30] This leaves open questions about how effective deliberation is in regulating the system's behavior in any given case. Just as we can imagine countries with ineffective governments, we can imagine agents with a deliberative standpoint whose behavior is in fact governed by whatever drive or appetite has momentary ascendance, wantons who act without overarching purpose.[31]

3.4. Diachronic Unity

What goes for the proper subject of a collection of synchronic judgments under a unifying standpoint one could also say goes also for the diachronic case. The "I" of "I remember," like the "I" of "I perceive," is a point of view that spans a collection of perspectives and is a formal product of the process that integrates them. This is the job of autobiographical memory that is charged with integrating the snapshot-like contents of episodic memory by providing those memories with a timeline and molding them into a consistent narrative. Locke held such a view. In his words,

> Consciousness, as far as ever it can be extended—should it be to ages past—unites existences and actions very remote in time into the same person, as well as it does the existences and actions of the immediately preceding moment: so that whatever has the consciousness of present and past actions, is the same person to whom they both belong.[32]

How can a cognizer that encounters an M at t and then another M at t' and then another at t'' recognize that there are three M's? Part of the answer is that he can combine the information in these states in a single further representation, that relates the first M encountered to the second and third. It's the same capacity that allows him to recognize an N encountered through one modality as the same N as that encountered through another. This process of combining and integrating wherein the questions about the sameness and difference of the things encountered in different perceptual episodes is synthesis. The proper subject of all of these thoughts—the "I" to whom they are all jointly ascribed—is the product of that activity. The nice thing about this view of the unity of the self is that it doesn't require the recognition of selves as primitive constituents of the universe. Constructing a point of view that spans modalities or spans temporal perspectives on this sort of view at the same time constitutes a self as occupant of that point of view.[33]

4. Echoes of Kant

I started out exploring the idea that we could find some clues about the identity of the subject by looking not at self-organizing systems but at countries and companies, that is, collectives organized into self-governing units. Someone else to whom this sort of analogy came naturally was Kant.[34] What Kant says about the self is part of an enormously complex tapestry of ideas, and (acknowledging some exegetical violence) I'm going to pull out two crucial insights and relate them to what preceded.

The first comes from Kant's discussion of the synthetic unity of apperception. And the second comes from his account of the formal subject of "I"-thoughts.

4.1. Synthetic Unity and the Perceptual Subject

First, to the first insight. On Kant's view, all thought is devoted directly or indirectly to the original function of thinking, whose role it is to (in his words, as they are usually translated) "combine and compare" objects of perceptual experience, and to recognize them under concepts. This process of "combining and comparing" is a subpersonal, preconscious activity that involves the integration of sensory information. Experience is the product of this activity and the world presents itself in experience as a spatio-temporally ordered whole. The separation of the sensory state from the object of perception and their relation to one another are both given in the content of the experience. The sensory state is assigned to a material presence in the landscape whose movement is tracked along with the changing landscape. Space and time, as they function in (this part) of Kant's discussion are just the forms, or ordering principles of the manifold that results when the sensory fields are integrated. Objects that occupy the manifold have properties that span the modalities. An object seen can also be felt, tasted, and smelled. Every sensory event that can be related to a location in that manifold is thereby related to every other. The sound of a voice can be related to a location in space and guide one's movements toward its source, and so on. When Kant talks about the synthetic unity of apperception, he is referring to this synthesis that integrates sensory information across the modalities.[35] When he writes in the B deduction—"All representations given to me must stand under this [unity], however they must be brought under it through a synthesis" (B 136)—I think he can be read as signaling both that experience is given to the subject as a synthetically unified whole, ↳ and—more important for our purposes—that both the subject and his experience are the product of synthetic activity.

4.2. Judgment: The Intentional Subject

Moving to the second insight, most commentators agree that Kant's analysis of the unity of "I" in "I think" is the most original aspect of his challenge to the Cartesian Cogito. The bulk of that analysis is given not in the Paralogisms of Pure Reason, but in the Transcendental Deduction of the Categories. There Kant characterizes the unity of the "I," in "I think," as a "formal" unity, by which he means that its role is not to refer to a thing whose identity provides the principle of unity for the states assigned to it but rather to impose standards of rational consistency on those states. Beatrice Longuenesse writes: "In [Kant's] view, 'I' is just the term to which we refer our thoughts in order to think of them as unified by one standpoint and bound by rules that commit us to bring about unity and consistency under a rational standpoint."[36]

This brings his view quite close to the suggestion I made that the transformation of beliefs into an inferentially articulated form in which they are explicitly self-ascribed and the implementation of procedures designed to check them for logical consistency and closure, and to make sure they meet the demands of rationality, is at the same time the construction of a rational subject to whom those states can be jointly ascribed.

It sounds very much more mysterious than it needs to in Kant, and I think the analogy with social collectives is genuinely helpful. Consider again the process by which the collection of opinions in a group of jurors is transformed into the collective opinion of the jury. All and only pronouncements in the collective voice are properly ascribed to the jury itself; the back and forth out of which the collective opinion is wrought is all "subjuridical." They belong to the jury members, but not to the jury. Rational constraints apply only to the opinions collected under a single standpoint, to the opinions self-ascribed in the collective voice. A pair of jurors can disagree with one another while remaining personally perfectly consistent, but the opinions of the jury pronounced in the collective voice do have to form a logically consistent set on pain of inconsistency. A pair of Supreme Court justices can hold differing opinions, but the pronouncements of the court made in the collective voice have to form a consistent set. The requirement of consistency is self-imposed by the self-regulating procedures we use to examine our commitments at both the personal and

institutional levels, and is reinforced by the social and institutional settings in which systems like persons, juries, and nations function.

Central here, again, are the two insights (1) that forging a point of view that spans perspectives gives rise to a subject that occupies that point of view, and (2) that the explicit self-ascription of representational states in an inferentially articulated form and implementation of procedures designed to achieve logical consistency and closure, is a condition of rational cognition.[37] The idea here, again, is that we have an "I" of "I see/hear/smell/taste . . ." as soon as we have a perspective that spans the modality. We have the "I" of "I think . . ." once we have the self-regulating procedures in place characteristic of rational cognition. We have the "I" of "I do" in place, once we have the "I" of "I think . . ." regulating action.

Now we are in a position to see our way around the two puzzles we started with: (1) why we don't see the self when we look at the brain, and (2) how anything material can exhibit the kind of unity that Descartes thought he apprehended in himself when he turned his gaze inward. The "I" of "I will/decide/affirm/judge/and do such and such," the one that makes all-things-considered judgments with an eye to the past and future, in the interests of the human being as a whole is the "I" of the reflective standpoint. It is the product of the distributed activity in the brain, not a neuron or synapse, a collection of neurons or synapses, or a homunculus lurking among them. It is not something one should expect to see in a close-up view of the microstructure of the brain, any more than one should expect to see the corporation walking the halls of company headquarters.[38] And we have an interpretation of the simplicity of the self. The agent, on this view, is not a material particular inside the brain, but a point of view (or perhaps the brain-body system unified by the point of view), whose indivisibility is a formal indivisibility, fully compatible with the idea that it is built out of material parts.[39]

5. Softening the Lines

Some qualifications are needed to make the self-governing structure less sharp around the edges. The claim was that when I say "I did such and such," in a sense that is explicitly meant to draw attention to the distinction between action and mere behavior, it is by reference to the role of the deliberative standpoint. But the deliberative standpoint plays a variety of roles in the determination of many different kinds of behavior. To make the account more general and more recognizable as a model of the human psyche, the dichotomy between action and mere behavior needs to be replaced by a nuanced understanding of the ↳ differences among roles. Eating, breathing, typing, driving, or maneuvering my legs across a tricky terrain, playing basketball, giving a lecture, and arguing with a colleague are all things I do, more or less attentively and more or less deliberately. The role that the deliberative perspective plays in these activities is as complex and varied as the role that the governing board of a company plays in managing its day-to-day affairs. And the deliberative process itself need not take the form of an explicitly articulated decision procedure. Even one's best all-things-considered judgments are often highly intuitive and nonprocedural. We have to acknowledge, moreover, that the reflective standpoint itself is a complex and evolving thing with its own internal disunities. How unified it is is a very personal matter that varies across the population and over time in the course of any one person's history. And the same can be said about how effectively it regulates behavior. Self-government comes in many different forms from the gentle guidance of a department manager who wants to be called in only for big decisions to the iron-handed control of a dictator who wants every decision to pass through her.[40] However these complexities get sorted out to give an accurate picture of the psychic life of the individual human being (something that moral psychologists have subjected to close and ongoing scrutiny), the self-governing structure, I suggest, provides the setting in which they are to be understood.

6. Conclusion

I started with the question, What kind of thing is the I that acts when I act, doing what I do? Is it the kind of thing that the neuroscientist should expect to see when he looks at the brain and can it have the kind of unity that Descartes noted when he turned his thought reflexively on himself? I argued that the self-governing structure introduces the agent into the picture without reifying it, giving us a non-question-begging, non-humuncular account of the "I" of "I do" (i.e., the proper subject of judgment and volition and the locus of responsibility). If you are looking for something in nature that has the right shape and plays the right role in the determination of behavior, then you shouldn't be looking for a concrete particular, or brain pearl lodged somewhere in the human machine. In a being with the kind of psychic organization exemplified by self-governing systems, the reflective standpoint is the thing that judges when I judge, decides when I decide, and is in control of the things that I do.

Acknowledgments

p. 292

Notes

1.   Dennett, 1991; Gazzaniga, 1992, 1998; Metzinger, 2003.

2.   Dennett, 1989, p. 163.

3.   Dennett, 1998, "Speaking for Ourselves," in *Brainchildren*, pp. 39–49.

4.   There are other arguments, of course, and those need to be addressed (see, for example, my "Saving the Baby," 2006), but this one is quite interesting, not only because it has some rhetorical force, but because it is instructive to see what is wrong with it.

5.   Descartes, (1993[1641]), p. 196.

6.   In linguistic mode: what does the "I" of "I did x" refer to?

7.   These are rough, heuristic divisions, not intended to be either exclusive or exhaustive.

8.   By "motion," I mean any variation in relative state.

9.   I follow my "Self-Organization and Self-Governance" (2011).

10.   We know that the collective good is not in general reducible to the good of the components of a system, even under very loose requirements on reducibility. There are special circumstances in which the collective good emerges from the self-interested behavior of components without any deliberate collectivization of practical reason, as in a free-market economy in a Smithian world. But that is the exception rather than the rule.

11.   "The will of the people" is ill defined until we have an algorithm that tells us how to compute it from aggregated opinion (i.e., until we have a function that defines the collective will as some weighted aggregate or function of individual opinions).

12.   In evolutionary terms, the collectivization of knowledge subserves the collectivization of deliberation. In the right kind of environment, for systems with complex enough interests, pooling information will produce more effective action. This does not mean that the more information the collective has at its disposal, the more effectively it can act. When there are

processing and attention constraints, a filter that lets only relevant information through might lead to more effective action.

13. See Rodney Brooks, *Cambrian Intelligence* (1999) and Andy Clark, *Being There* (1998), and more recently Anthony Chemero, *Radical Embodied Cognitive Science* (2011); see Kevin Kelly, *Out of Control* (1995) for an especially engaging account.

14. Just as what makes the covert activities of the Central Intelligence Agency (CIA) things that our government does, even if they go on largely outside the public eye and aren't passed by the legislature.

15. The collective voice can have both doxastic and volitional aspects. Referenda can determine public opinion, and elections give voice to the public will.

16. Individual members may share the opinions of the group, just as you and I may share opinions, but the opinion of the group is distinct from the opinion of any one or subset of its members.

17. I'm staying neutral about just what this "commitment to meeting the demands of rationality" involves; whatever norms govern doxastic and practical commitments on the personal level apply on the collective level for a system with a collective voice.

18. See especially Brandom, *Making It Explicit* (1994).

19. On this account of affirmation, see my "Immunity to Error" (2012) and "Decision and the Open Future" (2011).

20. Or whatever form such pronouncements take.

21. Not all first-personal intentional ascriptions are avowals. To get the right account of self-knowledge, we need a two-tier account along the lines of McGeer and Petit (2002), which allows for both descriptive and performative aspects of self-ascription.

22. This sort of avowal-based account of self-ascription of mental states has recently come into vogue in discussions of self-knowledge. See Gertler (2011) for an excellent discussion. Those accounts suffer from not distinguishing descriptive cases from cases involving judgment. There is good motivation for a hybrid account. On the distinction between credence and judgment, see McGeer and Petit (2002).

23. Integration of the informational streams leading from sensory surfaces is something the brain does for us. But it is not just these streams that need to be unified. It's also the many voices of past selves, each with its own constellation of conative and doxastic commitments. These selves are integrated only by self-conscious discipline and work.

24. This leaves us with some (tolerable) vagueness. What exactly are the contributors to my voice? The informational streams that it unifies originate in the environment and the distinction between body and environment is soft in ways emphasized in Dennett (1997) and Clark (1998). We can say the same thing about populations. Who, exactly, is included in the "we" of the people? The informational streams that get resolved into the collective voice are attenuated and there is no obvious terminus. There are the voters, of course, but also the people that voters talk to, and the news media and informal channels through which they get information, and so on. And then there is the question of the collection of parts that are controlled by the government; there are the citizens, of course, the polis, or body politic, but there is a much wider circle of influence affected by more attenuated links and only arbitrary or "legislative" boundaries. All of these sources of vagueness are present, though largely unacknowledged, in the case of the self.

25. We need to expand our vocabulary for the relations between parts and wholes in a manner that makes room for collectives in addition to collections and that interfaces with our criteria for individuating objects. There is no uniform usage here. ↳ Do we want to say you have new objects at the higher level? Or new agents but no new objects? Or is there some other way of describing these cases?

26. The "self" in self-regarding indicates the reflexive nature of the regarding relation. These are attitudes that take internal states as their object.

27. See my "Self-Organization and Self-Governance" (2011).

28. Of course that status can be revoked; entitlements are hostage to fulfillment of commitments. And nothing has been said

about what it takes to acquire the status.

29. One might wonder about the role of language and the social setting. Is the social setting indispensable? And could there be a nonlinguistic form of affirmation? Does the internal voice speak necessarily in a language that also has a public use? These questions are linked. It may be that affirmation, in the sense in which it figures in this chapter, is an activity that can only take place in a social setting. Assertions in a public voice carry cognitive and practical commitments that get their meaning in the social setting, and it may be that this public space is the source of normative notions like truth, accuracy, and warrant.

30. Can you have a point of view without an occupant of the point of view? When you have the point of view, you get the occupant for free. The transition from "there is a point of view whose contents span" to "there is a subject whose point of view that is" is an ontologically trivial grammatical transformation. Something of cognitive and practical interest happens when a complex system collectivizes epistemic and practical effort. The transition organizes old materials in a new way, but it doesn't introduce anything nonmaterial into the world.

31. Frankfurt (1971).

32. Locke, *Enquiry Concerning Human Understanding*, ch. 27, "Of Identity and Diversity." The Lockean view is by far the dominant view among naturalistic philosophers.

33. There is an interesting question here about whether the indivisibility that Descartes noted when he considered himself as subject of thought extends to the self as temporally extended subject of autobiography. Lockeans allow that persons have temporal parts, though there has been a great deal of dispute about how much disunity there is along the temporal dimension in a life. See, for example, John Christman (2004) and Galen Strawson (2004).

34. It also came naturally to Plato. See Plato, *Complete Works* (1997), Grube (1958), Vlastos (1971), Bobonich (1994), Burnet (1916), Cooper (1977), Hall (1974), and Joseph (1935).

35. Think of how we integrate visual information obtained from different spatial perspectives by mapping them into a representation of space that's invariant under transformations between perspectives. In formal terms, synthesis is a generalization of this process. We integrate information obtained through different sensory pathways by mapping them into a common frame of reference—in this case, a conception of the physical world conceived as the common source visual, tactual, and other experience.

p. 295 36. "Kant's "I think" versus Descartes's "I am a thing that thinks,"
http://philosophy.fas.nyu.edu/docs/IO/2575/longuenesse1.pdfp. 16.

37. Pat Kitcher writes: "The unity of self-consciousness is brought about through combination. In constructing representations of objects out of perceptual materials, the understanding also partially constructs a rational subject. It follows that, if perceptions are not such as to permit rational cognition, then it is not just that human beings would not be aware of themselves as continuing cognizers; their states would lack the relations of epistemic dependency characteristic of rational cognizers. They would lack the unity of a cognizer" ("The Unity of Kant's Active Thinker," ms., p. 18).

38. There is a choice we have to make when we detach the "self" from self-governed. It can refer to the whole organized system of components or the deliberative point of view that is the formal subject of intention and volition (the "I" of "I think" and "I do") and plays the role of the self in separating action from mere behavior. The analogous question in these other examples is that of whether the "We" of "We, the jury" refers to the collection of jurors or something more abstract, the deliberative point of view forged when they iron out a consensus. Common usage provides arguments on both sides and it matters which examples you consider. When we use the "we" in "we, the people" or "we, the jury," we tend to think of it as including all the jury members. But with corporate entities like British Petroleum, we tend to think that the thing that acts when BP acts or speaks when BP issues a statement is the collective standpoint. The people who work for BP don't identify so strongly with the collective standpoint (in many cases, there is a clear divergence between personal and corporate attitudes) and it, and in those cases, it is much more natural to think of the deliberative standpoint of BP as the agent (the thing that acts when BP acts).

39. The rights and responsibilities that come with self-governance will depend on what kind of self-governance we are talking about. There will be a good deal that is specific to say about political, economic, and corporate self-governance, but I suspect little that is general. There may be good arguments that the rights and responsibilities of individual human selves (whatever those turn out to be) place limits on those of corporate or collective selves.

40.   See Clark (2006) and Ismael (1997).

References

Bobonich, C. (1994). "Akrasia and Agency in Plato's Laws and Republic," *Archiv fur Geschichte der Philosophie,* 76, 3–36. 10.1515/agph.1994.76.1.3
Google Scholar   WorldCat   Crossref

Brandom, R. (1994). *Making It Explicit: Reasoning, Representing, and Discursive Commitment*. Cambridge, MA: Harvard University Press.
Google Scholar   Google Preview   WorldCat   COPAC

Brooks, R. (1999). *Cambrian Intelligence: The Early History of the New AI*. Cambridge, MA: MIT Press.
Google Scholar   Google Preview   WorldCat   COPAC

p. 296   Burnet, J. (1916). "The Socratic Doctrine of the Soul," *Proceedings of the British Academy,* 7, 235–259.
Google Scholar   WorldCat

Chemero, A. (2011). *Radical Embodied Cognitive Science*. Cambridge, MA: MIT Press.
Google Scholar   Google Preview   WorldCat   COPAC

Christman, J. (2004). "Narrative Unity as a Condition of Personhood," *Metaphilosophy,* 35(5), 695–713. 10.1111/j.1467-9973.2004.00345.x
Google Scholar   WorldCat   Crossref

Clark, A. (1998). *Being There: Putting Brain, Body, and World Together Again*. Cambridge, MA: MIT Press.
Google Scholar   Google Preview   WorldCat   COPAC

Clark, A. (2006). "Soft Selves and Ecological Control," in D. Spurrett, D. Ross, H. Kincaid, and L. Stephens, eds., *Distributed Cognition and the Will*. Cambridge, MA: MIT Press.
Google Scholar   Google Preview   WorldCat   COPAC

Cooper, J. (1977). "The Psychology of Justice in Plato," *American Philosophical Quarterly,* 14(2), 151–157.
Google Scholar   WorldCat

Dennett, D. (1989). "The Origins of Selves," Cogito, 3 (Autumn 1989), 163–173. Reprinted in Daniel Kolak and R. Martin, eds., *Self and Identity: Contemporary Philosophical Issues*. New York: Macmillan, 1991, p. 163.
Google Scholar   Google Preview   WorldCat   COPAC

Dennett, D. (1991). *Consciousness Explained*. New York: Little, Brown.
Google Scholar   Google Preview   WorldCat   COPAC

Dennett, D. (1997). *Kinds of Minds*. New York: Basic Books.
Google Scholar   Google Preview   WorldCat   COPAC

Dennett, D. (1998). *Brainchildren*. Cambridge, MA. MIT Press.
Google Scholar   Google Preview   WorldCat   COPAC

Descartes, R. (1993[1641]). *Meditations on First Philosophy*, Stanley Tweyman, ed., p. 196. London: Routledge.
Google Scholar   Google Preview   WorldCat   COPAC

Frankfurt, H. (1871). "Freedom of the Will and the Concept of a Person," *Journal of Philosophy,* 68(1) (January 14), 5–20. 10.2307/2024717
Google Scholar   WorldCat   Crossref

Gazzinga, M. (1992). *Nature's Mind*. New York: Basic Books.

Google Scholar   Google Preview   WorldCat   COPAC

Gazzaniga, M. (1998). *The Mind's Past*. New York: Basic Books.
Google Scholar   Google Preview   WorldCat   COPAC

Gertler, B. (2011). "Self-Knowledge," in Edward N. Zalta, ed., *The Stanford Encyclopedia of Philosophy* (Spring 2011 Edition), http://plato.stanford.edu/archives/spr2011/entries/self-knowledge/.
Google Scholar   Google Preview   WorldCat   COPAC

Grube, G. M. A. (1958). *Plato's Thought*. Boston: Beacon Press.
Google Scholar   Google Preview   WorldCat   COPAC

Hall, R. (1974). "Plato's Political Analogy: Fallacy or Analogy?" *Journal of the History of Philosophy,* 12, 419–435.
Google Scholar   WorldCat

Ismael, J. (1997). *The Situated Self*. New York: Oxford University Press.
Google Scholar   Google Preview   WorldCat   COPAC

Ismael, J. (2006). "Saving the Baby: Dennett on Autobiography," *Philosophical Psychology,* June 2006. 10.1080/09515080600690565
Google Scholar   WorldCat   Crossref

Ismael, J. (2011). "Decision and the Open Future," in Adrian Bardon, ed, *The Future of the Philosophy of Time*. New York: Routledge.
Google Scholar   Google Preview   WorldCat   COPAC

Ismael, J. (2011). "Self-Organization and Self-Governance," *Philosophy of the Social Sciences,* S 2011, 41(3), 327–351. 10.1177/0048393110363435
Google Scholar   WorldCat   Crossref

Ismael, J. (2012). "Immunity to Error as an Artefact of Transition between Representational Media," in Simon Prosser and François Recanati, eds., *Immunity to Error through Misidentification: New Essays*. Cambridge: Cambridge University Press. 10.1017/CBO9781139043274
Google Scholar   Google Preview   WorldCat   COPAC   Crossref

p. 297   Joseph, H. W. B. (1935). "Plato's *Republic*: The Nature of the Soul," in H. W. B. Joseph, *Essays in Ancient and Modern Philosophy*. Oxford: Oxford University Press.
Google Scholar   Google Preview   WorldCat   COPAC

Kelly, K. (1995). *Out of Control: The New Biology of Machines, Social Systems, and the Economic World*. New York: Basic Books.
Google Scholar   Google Preview   WorldCat   COPAC

Locke, John. (First published 1690). *Enquiry Concerning Human Understanding*, ch. 27, "Of Identity and Diversity."
Google Scholar   Google Preview   WorldCat   COPAC

McGeer, V. and P. Pettit. (2002). "The Self-Regulating Mind," *Language and Communication* 22, 281–299. 10.1016/S0271-5309(02)00008-3
Google Scholar   WorldCat   Crossref

Metzinger, T. (2003). *Being No One*. Cambridge, MA: MIT Press.
Google Scholar   Google Preview   WorldCat   COPAC

Plato. (1997). *Complete Works*, J. Cooper, ed.. Indianapolis: Hackett.
Google Scholar   Google Preview   WorldCat   COPAC

Strawson, G. (2004) "Against Narrativity," *Ration* 17 (4), 428–452. 10.1111/j.1467-9329.2004.00264.x

Google Scholar   WorldCat   Crossref

Vlastos, G., ed. (1971). *Plato II: Ethics, Politics, and Philosophy of Art and Religion*. New York: Doubleday.

Google Scholar   Google Preview   WorldCat   COPAC