UC Merced

Proceedings of the Annual Meeting of the Cognitive Science Society

Title

Cognitive Science and Two Images of the Person

Permalink

https://escholarship.org/uc/item/9sj3k7qx

Journal

Proceedings of the Annual Meeting of the Cognitive Science Society, 17(0)

Author

Brook, Andrew

Publication Date

1995

Peer reviewed

Cognitive Science and Two Images of the Person

Andrew Brook
Carleton University
Ottawa, ON., Canada K1S 5B6
University of Oxford
Oxford, England OX1 4AW
abrook@ccs.carleton.ca

Abstract

A certain indecisiveness and lack of common purpose seems to be a feature of cognitive science at the moment. We are in this paper that it can be explained in part by cognitive science's lack of success so far in connecting its scientific, computational image (better, images) of cognition to what we experience of people in ordinary life: in society, law, literature, etc. Following Sellars (1963), we call these two ways of representing cognizers the scientific image and the manifest image. The scientific image sees persons, and also artificial cognitive systems, as vast assemblages of postulated units of some kind. In the manifest image by contrast, persons are seen as unified centre of representation, deliberation and action, able to reach focused, unified decisions and take focused, unified actions. Since the manifest image is the murkier of the two, more of the paper is devoted to it than to the scientific image. The manifest image is richer and more diverse than might at first be thought.

Key words: philosophy/foundations; explanation.

Cognitive science has had less impact outside its boundaries than might have been expected given the huge range of discoveries it has made and the considerable resources expended on research into cognition in the past few decades. As a number of speakers at last year's workshop on Cognitive Science Education observed, cognitive science is still far from having a unified research programme and it is now forty years old. (Compare the neurosciences. Though much newer as a self-identified activity, their major international organization has over ten times the members of the Cognitive Science Society and its members don't spend time writing papers like this.) Even people within cognitive science often do not know how to take or to assess many of its claims. Part of the reason for all this, I think, is that so far the findings of cognitive science have not made much connection to persons as we experience and conceive of them in everyday life. An example.

In the last decades, we have developed many marvellous inference and decision systems. Now ask, what do these systems tell us about the activity of me and other people making a decision: the effort to think the situation through and reach a decision, then the effort (the same or a different kind of effort?) to adhere to the decision in the face of some unforeseen complication or -- a different kind of case -- against a strong desire to do otherwise? Very little; the operating principles of inference systems, etc., make almost no contact with decision-making as we find it in persons.

The point can be generalized. So far, cognitive science has had very little impact on social policy, interpersonal relationships, rules and structures for regulating interpersonal practice, the law, and so on. Where it has had an impact, moreover, it has generally not in the way we might expect, by telling us important things about what social actors are like, but as new and powerful prosthetics, in something like the way word processors have had an impact on writing. (Education is a partial exception, and one could think of others.) One reason, I think, is that most people have no idea how to apply the various conceptions of the human cognizer of cognitive science to the human person, no idea how to use them to understand themselves or others. Note that the problem is not just to do with introspection. How to apply cognitive science's notions of a cognitive system to others is as much a mystery as how to apply it to ourselves.

To forestall a possible misunderstanding, let me make clear that I am not setting the stage for some new form of mysterianism. (Flanagan calls philosophers like Nagel and McGinn who urge that consciousness, subjectivity or whatever is somehow beyond the ken of rational modelling New Mysterians.) To the contrary, if there is a problem here, it is a problem, something to be solved, not a mystery to be venerated. Veneration is not one of my favourite attitudes. Some philosophers have taken up aspects of the issue, if perhaps not the issue as a whole, Charles Taylor in particular.

Sellars' Two Images

So we seem to have two conceptions of persons and no good way to connect them. Sellars took a large first step towards getting these two conceptions clearer thirty years ago, in "Philosophy and the scientific image of man" (1963). He called the two conceptions *the manifest image* and *the scientific image.* The manifest image is the image of the human person of ordinary moral, social, and interpersonal life. The scientific image is the image of the person as a postulated system of smaller units of some kind, the image we find in cognitive science. The leading theory in this image nowadays is that persons are a vast assemblage of neurons tied together in complex biochemical and informational relationships. Sellars treats them in concert with a similar pair of images of the rest of the natural world but we will consider them in isolation here; he urges that the two images are radically different from one another.

'The' scientific image is in fact more than one image. A number of scientific images of the person have had an effect on our conception of persons, including: the neurological picture, the computational information-processing picture, and more recently the connectionist, distributed-representation picture. If we extend our considerations to include historical ones, the number is even larger; four of the more important were

Descartes' neural hydraulics, the neuron and its quantities of energy of Freud's *Project for a Scientific Psychology*, behaviourism's stimulus-response model, and the condition-action picture of classical production systems. However different these various pictures may be among themselves, they are also quite similar in one way: they all postulate an unobservable unit, much smaller than the person, out of which persons are built. Since they all differ from the manifest image in this respect, no harm is done by talk of `the' scientific image. (Sellars views behaviourism as a manifest-image theory. Since I think the manifest should be largely confined to the intentional, a point to which we will return, I think he is wrong about this, but it is a relatively minor point.)

The scientific and the manifest images are structured quite differently. In the manifest image, the basic unit is the whole person. The building block in the manifest image is nothing less than the human person, a being that can observe, make decisions, identify itself with things, enter into relationships with others, govern itself by standards, and so on. Taking a whole person as the basic unit is characteristic of much of social science and virtually all of practical, interpersonal or social activity. Both start from persons as a unit and focus on what moves these units, how they relate to other such units, how they relate to things in the nonpersonal world, and so on. More generally, the manifest image starts with the person as manifest in everyday life. The scientific image by contrast takes as its basic building block unobservable entities and unobservable processes of one kind or another, both postulated to *explain* features of the manifest (p. 19). The assumption is that the large basic unit of the manifest image consists of assemblies of the smaller basic unit of the scientific image, usually vast numbers of them. The two images just delineated are images of the human person as a whole, of all its essential aspects.

As Sellars conceives it, there is more to the manifest image than just taking the whole person as the basic unit of investigation. For Sellars, the manifest image of something is simply everything contained in the way we think about that kind of thing in everyday life. Indeed, it is more than this; Sellars also includes anything we could discover about the thing without postulating simpler, unobservable components as units out of which it is composed. With persons, this works out in the following way. We have been able to discover a great deal about what is correlated with what in people. Anger is correlated with insults, contentment with good relationships, adult disturbance with childhood traumata, etc. Many of these discoveries go well beyond what was contained in our original commonsense conception of the person. The manifest image we now have of persons thus includes not only the way it presents itself to us in everyday life but also everything we have been able to discover about persons by studying patterns of correlations.

To be sure, Sellars does not always honour the richness of his own conception. In particular, sometimes he seems to equate the manifest with the observable (p. 19). Even if we include the introspectible in the observable as Sellars does (p. 19), this characterization is too restrictive. There is a great deal in our manifest image of persons that is not observable or introspectible, very straight-forward things like character and levels of ability, for example. Neither is observable or introspectible, yet both are clearly aspects of our manifest image of the human person.

So what does characterize this image? Taking up a hint from the last paragraph, we might think that the manifest image is the arena of psychological explanation in the language of intentionality. The scientific image would then be characterized as the arena of mechanistic, i.e., non-teleological explanation. The idea at the heart of psychological explanation is the familiar one that objects of perception, desire, fantasy, belief, memory, etc. can have meaning for people. As a result, people can think and feel and do things for reasons, not just as the result of mechanistic causes. Reasons in turn have intentionality; they are about something, have an object. Psychological explanations in terms of reasons are very different from standard mechanistic explanations and part of the difference between the two images is clearly captured by this difference, but there is more to the manifest image than this. Mechanistic explanations of a kind also play a role in the manifest image, namely patterns of correlation. They are not postulational, so they are manifest-image, yet they are clearly not intentional either. The difference between the kinds of explanation does not exhaust the difference between the two images.

Fortunately, Sellars has a better suggestion: the manifest image is the framework in which we encounter ourselves (p. 6). That is to say, it is the framework within which we experience, reflect on, relate to, and interact with ourselves and one another. This suggestion provides something that our previous explorations left out. The distinction between the two kinds of explanation was focused on psychological content and attitudes, and modes of explaining them. A large part of our manifest image of persons, however, is an image of the thing that has this content, takes up these attitudes -- it is an image of a subject and agent. Moreover, as represented in the image, the subject and agent has a very particular character. Our job is now to describe this part of the manifest image.

Features of the Manifest Image

Probably the most important feature of our conception of the person, the basic unit of the manifest image, is that persons are originators of action, where by `action' we mean not just behaviour but behaviour that is a result of the formation of an intention and the taking of a decision. On a first pass, this being seems to have three prominent features:

(1) Persons are subjects and agents -- centres of representation, imagination, reflection, and desire, originators of intentions and decisions.

(2) Persons can guide themselves by reasons, things that motivate and so provide reasons for and against a belief or feeling or course of action in such a way that action ensues only as a result of decision, not simply by the unfolding of causal influences.

(3) Among the most important motivators or reasons for action are emotions -- fears, feelings of affection, feelings of gratitude, resentments, hostilities and biologically-based desires -- hungers, lusts, feelings of discomfort, etc.

All three features are distinctive of our image of the human person and it seems likely that all of them are also essential to it. What exactly we take to be built into any of three is a bit mysterious, but here are some candidates.

1. Making an effort. Sometimes it is easy to reach a decision, but sometimes it takes effort -- effort to understand a situation, effort to figure out what to do, effort, sometimes, to resist temptation and keep to a decision ('I really should reread three more pages of this paper but I would so much like to quit for the night'), and even effort to overcome obstacles, both human and non-human. Moreover, sometimes, try as hard as we might, we do not manage to do what we decided to do. What is making an effort like? Are all the exertions of effort just sketched of one kind or a number of different kinds? What is the agent who makes these various efforts like? It is hard to say.

2. Unity of focus. Making choices requires something else that is central in our manifest image of the person. I will call it unity of focus. To see what it is like, start with the better known unity of consciousness (UC). One form of UC is a matter of being aware of a whole group of representations at the same time; another consists of being aware of oneself as the common subject of those representations. We can define it more formally as follows:

> The unity of consciousness (UC) =*df.* (i) a representing in which (ii) a number of representations and/or objects of representation are combined in such a way that to be aware of any of these representations is also to be aware of other representations as connected to it. (See Brook 1994, Ch. 3).

Clearly UC is an important part of the manifest image of persons. However, I think a kind of unity found on the volitional side is even more central to what we conceive a person to be, what I call *unity of focus.* We conceive of persons as beings able to focus their intentional resources on courses of action. They can focus on a number of considerations at the same time and weigh up their implications. They can focus on a number of alternative courses of action at the same time, and assess them against one another in the light of desires, moral beliefs, wishes for other people, etc. They can bring these considerations together to form an intention and choose a course of action. And they can focus their intentional resources on carrying out that course of action, against obstacles, conflicting desires, and so forth. The unity of focus involved in all these activities seems to be something more than just a unified 'field' of consciousness.

It has often been remarked that when we think of persons within the manifest image, there is a powerful inclination to think of them as being or containing a grand unified homunculus, a little being who is the centre of consciousness, agency, responsibility, and so on. It seems likely that what I am here calling the unity of focus is one of the phenomena that contributes most strongly to this inclination. We will return to the problem of the homunculus briefly below.

3. Reference to Self. Persons can refer to themselves using indicators that have semantic properties different from those found in virtually any other indicator. Shoemaker's term for this form of reference is *self-reference without identification,* which he characterizes in the following way:

> My use of the word 'I' as the subject of [statements such as 'I feel pain' or 'I see a canary'] is not due to my having identified as myself something [otherwise recognized] of which I know, or believe, or wish to say, that the predicate of my statement applies to it. (1968, p. 558)

That is to say, I am aware of myself, as myself, without inferring this from any other feature of myself. If so, that the referent is myself is something I know independently of knowing anything else. If so, I must be able to refer to myself as myself independently of 'noting any quality' in myself, as Kant put it (1781, A355). If so, finally, the first-person pronouns are semantically quite unusual.

4. Anticipating a future as one's own. I have a striking ability to imagine a future person as myself, to imagine me having the experiences and doing the actions that I suppose he or she will have or do, even when I suppose at the same time that I am connected to that person by all manner of unusual connections and/or lack of connections. Williams explored this phenomenon in a series of interesting thought-experiments in (1973), esp. Ch. 3.

All of (1) to (4) seem to be either parts of our manifest image of the person or natural extensions of it. Two things are striking about the resulting conception of the person. One is its central importance in practical and especially interpersonal and social life. The other is how little anybody in philosophy, cognitive science or anywhere else has managed to say about it. In particular, little progress has been made with capturing any of (1) to (4) in any kind of postulational, mechanistic model, whether neuronal, computational, connectionist or otherwise. And that is one reason, I would argue, why cognitive science has had less impact on activities that revolve around the manifest image of the person than might have been expected.

The same gap may explain something of the somewhat fragmented and scholastic state of the functionalist model of the mind. In much the same way as we do not know how to map manifest image phenomena onto any mechanistic model, we do nt know how to map manifest phenomena onto any of the models of the current flavours of functionalism.

Dennett and Fodor: Two Approaches to the Manifest Image

Dennett and Fodor between them have the two most prominent strategies for addressing the problem at the moment. It is helpful to consider their approaches in the context of something specific, so let us return to one of the features of the manifest image mentioned earlier about which more needs to be said, our inclination to think of persons in it as being or containing a homunculus at the centre of our representational world. It is likely that this picture probably underlies all the other aspects of the manifest image mentioned above.

Dennett's approach to the manifest image is to eliminate the problem by eliminating the image, at least as anything worth taking seriously. His book on consciousness (1991) is a most thorough-going attempt to undermine our conviction that something of substance corresponds to the manifest image. With respect to our specific inclination to think of ourselves and others in terms of a grand homunculus, Dennett attempts, via a variety of ingenious thought-experiments, to eliminate any urge to think that there might be something that accords with this inclination. His approach, over-simplified a bit, is to urge that the problem is our sense that there is a problem. There is no problem. We think there is only because we misinterpret 'phenomenology' in certain natural ways, phenomenology being the way we appear to ourselves and to others.

I do not think this move will take us far. Here is why. In my view, Dennett's approach merely shifts the problem. The puzzle now becomes the phenomenology itself: what could produce

such a thing as a sense that I am a homunculus? How could we have been led so astray -- how could a conception of such profound, even ineliminable social importance as our manifest conception of the person be at it core nothing more than a huge mistake? (This in outline is all I think most eliminative strategies achieve; they just shift the puzzle.)

By contrast, Fodor *is* willing to mechanize the manifest image, or try. His approach is to chip away at the edges: sort out a syntax module here, a vision module there, and postpone the rest to a better day. The problem with this approach is that, while the chips are eminently worth knocking off and we have learned a great deal about the abstract and in the case of vision even the neural structure of some encapsulated, nonconscious subsystems by doing so, the central person, the big homunculus, is not touched and remains as much a mystery as ever. All Fodor's and others' discoveries about the various nonconscious cognitive subsystems in a person have done is to shrink the range of the unified subject and agent, shrink the homunculi. They have done little or nothing either to discharge it or to tell us what it is like.

Fodor seems to be beginning to worry about this himself; his latest book (1994) contains a number of uneasy references to the centrality of consciousness and his inability to say anything informative about it.

Why do we need to `reduce' the homunculi? Why not adopt Dennett's strategy and just ignore it or explain it away? For this reason. On the one hand, we cannot dispense with it; we could as soon dispense with the language of love, law, psychology, motive, feeling, and representation. On the other hand, as the Churchlands and others have pointed out, the manifest image does not take us very far. In fact, it works at all only so long as the person being explaining is functioning well. Introduce any amount of cognitive or emotional damage or breakdown and the manifest image instantly `claws the air', as Paul Churchland puts it (1984, p. 46). And so on. Abandoning the manifest image is not one of our options (on this point I do not agree with the Churchlands) but abandoning the drive to capture it in mechanical models that illuminate and help us understand it is not an option either.

With respect to the two images, most researchers in cognitive science right now are either manifest image extenders or manifest image ignorers. There are few manifest image reducers around. Philosophers are prominent in the first group: they spend most of their time trying to find ways to capture manifest image phenomena in extensions and developments of the same kind. These folk worry hardly at all about how their ideas might be realized in a computational system.

The second group is made up of most of the rest of cognitive science -- the researchers who spend their time developing ever deeper and broader postulational models. These folk give little thought to how their models can capture manifest image phenomena. Sometimes one finds both sides in a single researcher, yet totally disconnected. I am thinking of cognitive psychologists I have known, for example, who work with almost nothing but a postulational model in their labs, but delight in playing around with notions such as the unity of consciousness outside it.

What is Missing?

What is missing on both sides is an effort *to bring the two images together*, to find postulational, computational accounts that illuminate phenomena of our everyday experience of ourselves and others. And that, I contend, is one reason why cognitive science has not made more progress. In the absence of robust roots in experience, we not only do not know how to assess a model, we do not even know how to take it. As Dennett once said on a related issue, clearly a meeting of minds is in order.

References

Brook, A. (1994). *Kant and the Mind.* Cambridge University Press.

Churchland, P. (1984). *Matter and Consciousness.* Bradford Books/MIT Press.

Dennett, D. (1991). *Consciousness Explained.* Little, Brown.

Flanagan, O. (1984). *The Science of the Mind.* Bradford Books/MIT Press.

Fodor, J. (1983). *The Modularity of Mind.* Bradford Books/MIT Press.

----------. (1994). *The Elm and the Expert.* Bradford Books/MIT Press.

Freud, S. (1895). Project for a scientific psychology. *Standard Edition of the Complete Psychological Works of Sigmund Freud,* trans and ed J. Strachey. Hogarth Press and Institute of Psycho-Analysis.

Kant, I. (1781). *Critique of Pure Reason.* Trans. Norman Kemp Smith as: *Immanuel Kant's Critique of Pure Reason.* Macmillan, 1963.

Sellars, W. (1963). Philosophy and the scientific image of man. In: *Science, Perception and Reality*, pp. 1-40. Routledge and Kegan Paul.

Shoemaker, S. (1968). Self-reference and self-awareness. *J. of Phil.* 65 no. 20 (1968), pp. 555-567.

Taylor, C. (1985). *Philosophical Papers*, 2 vols. Cambridge University Press.

Williams, B. (1973). *Problems of the Self.* Cambridge University Press, esp. Ch. 3, The self and the future.