BETWEEN
PERCEPTION
AND ACTION

BENCE NANAY
Between Perception and Action
To Felicitas
# Contents

_Acknowledgements_  ix

1. Introduction  1
   1.1 The icing and the cake  1
   1.2 Pragmatic representations  3
   1.3 Product differentiation  7

2. Pragmatic Representations  13
   2.1 Mediating between sensory input and motor output  13
   2.2 The immediate mental antecedents of actions  15
   2.3 The direction of fit  19
   2.4 Pragmatic representations are perceptual states  21
   2.5 The argument  23
   2.6 Potential counterexamples  28
   2.7 Pragmatic representations everywhere  31

3. Perception  33
   3.1 Are there perceptual representations at all?  33
   3.2 Perceptually attributed properties  36
   3.3 Sensory individuals  49
   3.4 Pragmatic representations and the dorsal visual subsystem  62

4. Action  67
   4.1 Belief–desire psychology and its discontents  67
   4.2 Naturalizing action theory  75
   4.3 Semi-actions  81
   4.4 The last refuge of the belief–desire model  86

5. Pragmatic Mental Imagery  102
   5.1 Mental imagery  102
   5.2 Mental imagery versus perception  105
   5.3 Pragmatic mental imagery  111
   5.4 Pretense  115
   5.5 Aliefs  121
This book is about perception, action, and what’s in between. It’s about the whole mind, in other words. It is written for philosophers, but not only for philosophers: it is also for psychologists or other empirical researchers of the mind with theoretical interests, as well as people who are just interested in understanding how the human mind works and how it is very similar to animal minds.

I started exploring the central claims of this book almost 20 years ago. As far as my philosophical attention deficit disorder allowed, I have been pursuing these claims ever since. In 1995, I published a very bad paper in a psychology journal on how we should reject Gibson’s views about perception while preserving the insight about the action-oriented nature of perception. Then I wrote a still pretty bad MPhil thesis at Cambridge University on the “special” mental states that mediate between sensory input and motor output, and then an only slightly better PhD at the University of California, Berkeley on action-oriented perception. I am grateful to my various teachers and colleagues throughout these years (and admire their patience) for not discouraging me from pursuing this project, given how unconvincing some of my arguments really were. The hope is that the arguments in this book are much more convincing.

Given that I have been talking to philosophers, psychologists, ethologists, and neuroscientists about the topic of this book for nearly 20 years, there is no way I can remember everyone who helped me with thinking more clearly about the topic of the book, or even provided written comments on earlier written-up versions of some of the ideas in the book. So I won’t even try—if I tried, I would, no doubt, leave out many people. But I am especially grateful to those who read the whole penultimate version of the manuscript and provided detailed feedback: Jonathan Cohen, two anonymous referees from Oxford University Press, and my postdocs, Maja Spener, Carolyn Dicey Jennings, Will Davies, and Craig French.

Various parts of the book were presented at the following symposia, conferences, and workshops: American Philosophical Association, Pacific Division (San Diego), Annual Conference of the Cognitive Science Society (Vancouver), American Philosophical Association,
Pacific Division (San Francisco), American Philosophical Association, Pacific Division (Vancouver), Annual Conference of the Cognitive Science Society (Berlin), Intuitions, Experiments and Philosophy Conference (Nottingham), Social Cognition Workshop (Cambridge), Evolution, Non-propositional Intentionality Workshop (Geneva), Co-operation and Rationality Conference (Bristol), Philosophy of Psychology, Neuroscience and Biology Conference (Oxford), Empathy/Sympathy Conference (Antwerp), Social Science Roundtable (Paris), Systematicity Workshop (San Jose), Phenomenal Qualities Conference (Hertfordshire), Animal Cognition Conference (Bochum), Philosophical Insights Conference (London), Philosophy of Science Association Biannual Meeting (San Diego), 5th Online Consciousness Conference, New Waves in Philosophy of Mind Online Conference.

I also gave talks based on the material in this book at the University of Bochum, University of Fribourg, Institut Jean Nicod, University of Edinburgh, Simon Fraser University, University of British Columbia, University of Cardiff, University of Turin, University of Geneva, Rice University, Syracuse University, York University, Concordia University, and the University of California, Berkeley. I am extremely grateful for the comments and grilling I received after these talks. I also learned a lot from teaching some of the material in the book—both at PhD seminars and at the undergraduate level—at Syracuse University, the University of British Columbia, and the University of Antwerp.

Finally, I am very grateful to the editors at Oxford University Press, especially Peter Momtchiloff, Eleanor Collins, Daniel Bourner, and Sarah Parker, for making the whole process as smooth as possible to Sally Evans-Darby for proofreading, and to Beth Morgan for compiling the index. The research was supported by the EU FP7 CIG grant PCIG09-GA-2011-293818 and the FWO Odysseus grant G.0020.12N.

Bence Nanay

Cambridge
February 19, 2013
1

Introduction

We are only philosophers when at work in our study: so soon as we indulge freely in the impressions of our senses, we are children like everyone else. Abstract speculation is a forced state of mind: our natural condition is that of everyman.

Jean-François Marmontel: Poétique Françoise, 1763.

1.1 The icing and the cake

There are two general approaches towards trying to understand the human mind. The first one is to take our sophisticated, highly intellectual, rational, linguistic, and uniquely human mental capacities to be paradigmatic, and try to understand how our minds work by focusing on them.

The second way of trying to understand the human mind emphasizes the continuity between human and animal minds. It therefore focuses on those simpler mental capacities that non-human animals also possess: those mental processes that make it possible for us to perform actions successfully—to put food in our mouth, or to get around without bumping into things.

As philosophers are in the business of giving rational, sophisticated, highly intellectual arguments, a tempting route for philosophers of mind is to project their own methodology onto the phenomenon they are trying to explain, and consider the human mind as primarily consisting of rational, highly intellectual, sophisticated mental capacities. I think we should resist this temptation.

My aim is to argue that the vast majority of what goes on in our mind is very similar to the simple mental processes of animals. Our complex, sophisticated, rational, and linguistic abilities could be described as the icing on the cake. Although the icing is what makes the cake look great,
we should try to understand the cake without worrying too much about the icing. In other words, the right methodology for philosophy of mind is to understand those simple mental capacities that we share with animals first, and then explain those uniquely human, highly intellectual mental capacities that make the human mind so remarkable. It is important that I am not suggesting that the human mind is to be understood as the animal mind plus some extra features—imagine instead a cake where the icing nicely seeps into the dough. But if we want to understand the human mind, we need to start with the simplest mental capacities and work our way towards the more complex ones.

To stretch the cake analogy even further, the aim of this book is to identify what could be described as the main ingredients of this cake. The basic units of our linguistic, higher order mental processes are conceptually/linguistically structured propositional attitudes: beliefs, desires, thoughts. But we have no reason to suppose that the basic units of those simple mental capacities that we share with animals would also be such propositional attitudes. In fact, the point could be made that when we describe the human mind as consisting of beliefs and desires, we are mirroring language. The basic units of language are sentences or utterances that express propositions; thus, it would be tempting to say that the basic units of our mind must be mental states that also express propositions: propositional attitudes. But if we maintain that our understanding of the human mind is not to be modeled on language or even on uniquely human linguistic thoughts, then we have no reason to accept that beliefs and desires are the basic building blocks of the human mind.

The need to look for new candidates to replace beliefs and desires as the basic building blocks of the human mind is not new. Gareth Evans, for example, suggested that we should consider information-carrying states to be the basic units instead (Evans 1982). I would like to make a more specific proposal. My claim is that the vast majority of what is going on in the human mind can be understood if we consider what I call “pragmatic representations” to be the basic units of our mental capacities.

The human mind, like the minds of non-human animals, has been selected for allowing us to perform actions successfully. And the vast majority of our actions, like the actions of non-human animals, could not be performed without perceptual guidance. My claim is that the mental state that mediates between sensory input and motor output is the basic
building block of the human mind. I call mental states of this kind pragmatic representations.

1.2 Pragmatic representations

What mediates between sensory input and motor output? This is probably the most basic question one can ask about the mind. There is stimulation on your retina, something happens in your skull, and then your hand reaches out to grab the apple in front of you. What is it that happens in between? What representations make it possible for you to grab this apple? My answer to these questions is that it is pragmatic representations that mediate between sensory input and motor output; it is a pragmatic representation that makes it possible for you to grab the apple.

Pragmatic representations are, at first approximation, the representational components of the immediate mental antecedents of action. They are also genuine perceptual states. The immediate mental antecedents of action are what make actions genuine actions. They constitute the difference between actions and mere bodily movements. They guide our ongoing bodily activities. And pragmatic representations are the representational components of these immediate mental antecedents of action.

These mental states represent the world as being a certain way: they are about the world, they refer to the world. In other words, they have representational content: they represent objects as having certain properties. This, however, does not mean that they must have a syntactically articulated propositional structure, or that they really are sentences written in some mental language. Pragmatic representations can be correct or incorrect. If they are correct, they guide our bodily activities well. If they are incorrect, they guide us badly.

What properties do pragmatic representations represent objects as having? Suppose that you want to pick up a cup. In order to perform this action, you need to represent the cup as having a certain spatial location, otherwise you would have no idea which direction to reach out towards. You also need to represent it as having a certain size, otherwise you could not approach it with the appropriate grip size. And you also need to represent it as having a certain weight, otherwise you would not
know what force you need to exert when lifting it. I call these properties “action-properties”: action-properties are properties that need to be represented in order for the agent to perform the action. Pragmatic representations attribute action-properties: they represent objects in an action-oriented manner (Nanay 2011a, 2011b, 2012a). And they typically attribute these action-properties unconsciously.

As pragmatic representations are supposed to play an important role in action, and as they are supposed to be perceptual states, if we accept that pragmatic representations are crucial for understanding our mental life then this will have radical consequences both for debates in philosophy of perception and for debates in philosophy of action.

Let us take philosophy of perception first. Pragmatic representations are bona fide perceptual states. It is an important question in the philosophy of perception what properties we perceive objects as having. Shape, size, and color are obvious candidates. It is much less clear whether we represent sortal properties perceptually: whether we literally see objects as chairs and tables and not just infer that they are. I argue that we sometimes also perceive objects as having action-properties.

This should not sound particularly surprising. Our (extremely complex) perceptual system was selected for helping us to perform actions on which our survival depended. It is hardly surprising, then, that it was selected for representing objects as having properties that are relevant to the performance of our actions: as having action-properties. As Tyler Burge says, “since perception guides action, it is not surprising that perceptual kinds mesh with [our] activities” (Burge 2010, p. 324).

Importantly, my claim is that we perceive objects as having action-properties some of the time, not that we always do so. Pragmatic representations are perceptual states, but not all perceptual states are pragmatic representations. Very often we perceive objects in a way that is not action-oriented—for example, when we are admiring the landscape, sitting on a bench, in a detached manner, without any urge to perform any action at all.

Now let us see the relevance of pragmatic representations for the philosophy of action. One of the most important questions of philosophy of action is this: what makes actions actions? How do actions differ from mere bodily movements? The difference is some kind of mental state that triggers, guides, or maybe accompanies the bodily movement. But what is this mental state?
I argue that these immediate mental antecedents of our actions are partly constituted by pragmatic representations. These pragmatic representations attribute properties, the representation of which is necessary for the performance of the action, and they do so perceptually: they perceptually guide our bodily movement. But if this is true, then the belief–desire model of motivation needs to be adjusted or maybe even discarded.

According to the “belief–desire model” of motivation, beliefs and desires are necessary for motivating us to act. Perception may play a role in bringing about the belief that our action is based on, but it does not play any more substantial role. Here is an example. I look out of the window and I see that it is raining outside. I form a belief that it is raining outside. I have a desire not to get wet and this, together with my further (instrumental) belief that the best way of not getting wet in rain is to take an umbrella, leads to the forming of an intention to take an umbrella, and this intention triggers my action of taking my umbrella (see, e.g., Smith 1987, Davidson 1980, the origins of the account may go back to David Hume). In this model, beliefs and desires mediate between sensory input and motor output. Although this model does describe the way in which some of our (highly intellectual and arguably uniquely human) actions come about, it only applies to a very small minority of the actions we perform. Some of our perceptual states, namely pragmatic representations, play a more crucial role in bringing about our actions than has been acknowledged.

In the case of most of the actions we perform, the only mental states (more precisely: the only representational states) that mediate between sensory input and motor output are pragmatic representations—and not a set of beliefs and desires, as the classical belief–desire model suggests. Further, even in those cases when beliefs and desires play an essential part in bringing about our actions, pragmatic representations still need to be involved (for example, when I actually pick up the umbrella).

We have seen two aspects of the philosophical relevance of pragmatic representations: each time we are performing an action, we have a pragmatic representation. And many of our perceptual states are pragmatic representations.

But there are further reasons to be interested in mental states of this kind. Animals and small children are capable of performing goal-directed actions, such as running away from predators or chasing prey.
But if they are, they must be able to have pragmatic representations. Hence, even organisms that may be incapable of entertaining complex thoughts and beliefs must be able to have pragmatic representations.

These mental states are both phylogenetically and ontogenetically quite basic. It can be argued that the first mental representations that appeared in the course of evolution were pragmatic representations. And, similarly, the first mental representations that appear in the course of child development are also likely to be pragmatic representations. If we want to explain the minds of non-human animals or small children, we would be well advised to focus on mental states of this kind.

The general rhetoric I am following here is to examine the simple mental processes that animals and humans are equally capable of, and postpone the analysis of uniquely human, complex, sophisticated mental capacities. But if we accept the general framework I propose, we can go a bit further and at least begin to explain some of our more complex mental capacities, such as our engagement with, and understanding of, others, and even basic emotions such as empathy.

As we have seen, pragmatic representations represent objects in a way that is relevant to my action. But there is an intriguing phenomenon in the vicinity: mental states that represent objects in a way that is relevant to someone else’s action. I argue that these mental states constitute our simplest and most basic capacity to engage with others. I call this way of engaging with others vicarious perception, and argue that a number of big debates—not only in philosophy, but also in cognitive ethology and developmental psychology—about our ability to attribute mental states to others (an ability usually called “theory of mind”) could be resolved if we restrict their scope to vicarious perception.

In the last three decades, the concept of “theory of mind” has been at the center of interest in philosophy of mind, psychology, and primatology. Some important questions about “theory of mind” are the following: Do non-human animals have theory of mind? How does theory of mind develop in ontogeny? What mental processes make theory of mind possible in humans? What are the neural underpinnings of theory of mind?

While these questions are difficult to tackle (and there has been no sign of any consensual answer) so long as they are about “theory of mind,” if we take them to be about vicarious perception (and not theory of mind) we get straightforward and nontrivial answers. More precisely,
it can be argued that all experiments that are supposed to show that non-human primates have theory of mind in fact demonstrate that they are capable of vicarious perception. The same goes for the experiments with infants less than 12 months old. If we shift the emphasis from theory of mind to vicarious perception, we can make real advances in understanding the origins of social cognition.

1.3 Product differentiation

The traditional way of thinking about the human mind has been to explain it using the analogy of symbol-manipulating serial computers. Sensory input comes in, it gets coded into some kind of symbolic, sentence-like propositional format and is matched against other symbols, and this sometimes gives rise to motor output (Fodor 1981, 1983, Davidson 1980). This is the picture the traditional belief–desire model uses, where propositional attitudes (beliefs, desires, intentions) mediate between the sensory input and the motor output. There is no agreement about just what kind of propositional attitudes these mediating states are—and it very much depends on one’s concept of propositions. What I take to be the traditional way of thinking about the human mind takes the mediators between sensory input and motor output to be propositional attitudes whose content is syntactically, conceptually, and maybe even linguistically articulated. I will call this view the “computationalist/propositionalist” account (see Figure 1). Those who take propositions to be necessarily syntactically, conceptually, and maybe even linguistically articulated could just read “propositionalist” here. I add the “computationalist” label because those who have a weaker notion of propositionality may object to calling this view propositionalist per se. Thus the label: “computationism/propositionalism.”

![Figure 1 Computationalism/propositionalism](image)
The most important alternative to the computationalist/positionalist account of the human mind could be labeled the anti-representationalist/enactivist approach. According to the enactivist, perception is the active exploration of one’s environment, which can be described without talking about representations at all (Noë 2004, Hurley 1998, Gibson 1979). The proponents of this approach reject over-intellectualizing the mind by shifting the emphasis from mental to bodily activities: by emphasizing the importance of bodily coping—skills and abilities that are not really (or not exclusively) mental—they tend to emphasize the embodied nature of cognition. And this tends to amount to denying that perception is supposed to be described in representational terms at all.

According to this anti-representationalist/enactivist view, “perception is not a process of constructing internal representations” (Noë 2004, p. 178, see also Ballard 1996, O’Regan 1992). The main enactivist claim is that we have all the information we need in order to get around the world out there, in the world. So we do not need to construct representations at all and, more specifically, we do not need perceptual representations. As Dana Ballard put it, “the world is the repository of the information needed to act. With respect to the observer, it is stored ‘out there’, and by implication not represented internally in some mental state that exists separately from the stimulus” (Ballard 1996, p. 111, see also Brooks 1991, Ramsey 2007, Chemero 2009, Hutto and Myin 2013). Sensory input and motor output are so closely intertwined in a dynamic process that we do not need to posit any representations that would mediate between them (see Figure 2).

---

**Figure 2** Anti-representationalism/enactivism
I think that this is a mistake, and that rejecting talk of mental representations altogether because of the justified mistrust of a specific kind of representation—namely, conceptually/linguistically structured propositional attitudes—is a bit like pouring out the baby with the bathwater. These kinds of propositional attitudes are not the only kind of representations. My view is that the mind is to be understood in terms of representations, but that these representations are not conceptually or linguistically structured, nor are they uniquely human. They are better compared to the mental representations of the predator that make it possible for it to catch its prey. These representations are perceptual representations and inherently action-oriented (see Figure 3).

![Figure 3 Pragmatic representations](image)

In short, the novelty of my proposal is to carve out an intermediary position between the enactive/antirepresentationalist and the computationalist/propositionalist accounts. This intermediary position is supposed to inherit the explanatory advantages of both extremes.

It inherits the general view that representations are the basic building blocks of our mind from the computationalist/propositionalist accounts. The general upshot is simple: we should talk about the human mind in terms of representations. So the general theoretical framework is the same as the one the computationalist uses: sensory inputs give rise to some representations, which give rise to motor outputs.

But these intermediary representations in my framework and those used in the computationalist framework are very different. While the computationalist takes these representations to be propositional attitudes (that follow, more or less, a belief–desire model of connecting sensory input to motor output), I argue instead that these representations are (a) perceptual representations and (b) inherently action-oriented.

The main advantage of my approach over the computationalist one is that I can talk about the mental representations of animals and small children (as well as the mental representations of adult humans that make most of our everyday actions possible) without