ANDY CLARK
AND HIS CRITICS

EDITED BY

MATTEO COLOMBO
ELIZABETH IRVINE
MOG STAPLETON

Andy Clark and His Critics

Andy Clark and His Critics

EDITED BY
MATTEO COLOMBO
ELIZABETH IRVINE
and
MOG STAPLETON

CONTENTS

FOREWORD

The ideal of "all things considered" is a perpetual motion machine for epistemology. You can't think effectively about any hard topic without relying on unexamined constraints, self-imposed barricades on your imagination that permit you to dismiss without a hearing most of the myriad candidate solutions to whatever problem you are concentrating on. The need for brusque heuristic pruning of the search tree is a fact of life, not just a fact of chess. My colleague Marcel Kinsbourne has proposed that what makes any problem hard is always the fact that something attractive, and false, stands in the way, securing allegiance that then poisons the investigation. Philosophers, at their best, specialize in "opening our eyes to new perspectives," helping us overcome our subliminal aversions and loosen our grip on "home truths" that are so familiar we never stop to consider them. This delicate task calls for an artful mingling of arresting observations, vivid language, and a deep understanding of the theoretical work that has created the arena in which current disputes play themselves out.

Andy Clark is a peerless perspective-shifter, and the fruits of his labors are magnificently on display in this volume. There is a palpable sense of intellectual community and progress on the tough issues, a sense that if we put our heads together, we can discern the contributions of a wide variety of apparently warring *positions*, both philosophical and scientific, and weave them into a cable of mutual agreement, defeasible of course, but a new path of common ground on which to pursue further research. The central cord of Andy's cable, still only reluctantly endorsed or even resisted by some of the contributors, is (in my hardly disinterested opinion) the working assumption that there is nothing like *élan vital*, or *wonder tissue*, or *intrinsic intentionality* that distinguishes the mind from the rest of nature. One manifestation of this assumption is the Parity Principle (and the Reverse Parity Principle proposed by Goldstone), and another is the fruitful pursuit by all participants of continuities between all evolved organisms and indeed all designed tools for thinking, from writing and prostheses to Scrabble tiles and computers.

We totter, as usual, on the shoulders of giants: Descartes, Bayes, Skinner, Gibson, Fodor, and others, wielding the multitude of *isms* provoked or inspired by them, and one thing that strikes me about the discussions here is how often they expose a pendulum swinging between overstatement and oversimplification. You pay a price for the vividness without which you cannot hold the attention of your students (or colleagues): the Heartbreak of Premature Definition. What is *your* definition of functionalism, cognitivism, enactivism, embodied cognition, representation, affordance? There are good reasons to postpone definition until after you've elaborated some of the treasures, prospects, and risks of your position, but if you decline to define your terms, your critics will be obliged to define them for you, and the result can degenerate into a food fight of counterexamples and *reductios*. Is functionalism really just behaviorism unleashed by Skinner's 1964 observation that "the skin is not that important as a boundary"? What does cognitivism add to behaviorism, if not a dread homunculus? Once you've laundered all the *intellectualist* connotations out of the concept of *representation*, dismantling the inner user, how do representations differ from resonant loops? Concentrating myopically on devising variant definitions, *chisholming* away in a fugue state of defensive strategizing, is a well-known philosophical foible, *but it is not much in evidence here!* Andy Clark sets a fine example, followed by his critics, of really trying to educe the best insights from the opposition, and both the essays and Andy's gracious response to them provide models of philosophical behavior that should inspire and instruct all who enter these arenas.

This is not your grandfather's philosophy of mind, and the contributors are the all-stars of the twenty-first century. Among the many volumes of essays exploring the ideas of one philosopher or another, this one has a rare virtue: as you read the essays *you keep learning new things*, not just novel arguments, novel objections, novel critiques, but *facts about the world outside philosophy* that philosophers ought to know. Who would have thought that the ingenious devices of insect mating competition, or the ethnography of ritual memory boards among the Luma, or the architecture of Differential Neural Computers, or the function of postural sway, or the effects of left-handedness would illuminate any philosophical controversy? As one who gets impatient with philosophical writing that does not inform me about anything beyond the cleverness of the author, I find rewards aplenty in these pages, and I cannot think of a book that better exposes the limitations of traditional, factually impoverished philosophical combat. And like Clark's own work, it is all eminently readable by nonspecialists. Too many of our colleagues in philosophy have been subliminally taught that philosophy that isn't hard to read is not worth reading. This cheerful, and cheer-inducing, book is a fine counterexample.

Daniel C. Dennett

ACKNOWLEDGEMENTS

CONTRIBUTORS

Fred Adams, Department of Linguistics & Cognitive Science, University of Delaware

Kenneth Aizawa, Department of Philosophy, Rutgers University, Newark

Michael Anderson, Department of Philosophy, Western University, Ontario

Louise Barrett, Department of Psychology, University of Lethbridge

David J. Chalmers, Department of Philosophy, New York University

Anthony Chemero, Departments of Philosophy and Psychology, University of Cincinnati

Matteo Colombo received his PhD from the University of Edinburgh in 2012, supervised by Andy Clark and Peggy Seriès. He is now an associate professor at the Tilburg Center for Logic, Ethics, and Philosophy of Science, and at the Department of Philosophy, Tilburg University. He's interested in questions about evidence and explanation in the computational, cognitive, and brain sciences, and more generally in how the scientific and manifest images of mind relate to one another.

Katalin Farkas, Department of Philosophy, Central European University

Karl J. Friston, Institute of Neurology, University College London

Robert L. Goldstone, Department of Psychological and Brain Science, Indiana University

Jakob Hohwy, Department of Philosophy, Monash University

Elizabeth Irvine received her PhD from the University of Edinburgh in 2011, supervised by Andy Clark, and held postdoctoral research positions at the University of Tübingen and the Australian National University before becoming a lecturer at

Cardiff University. Her research interests are primarily in philosophy of cognitive science and psychology and philosophy of science.

David Kirsh, Department of Cognitive Science, University of California, San Diego

Geoff Lee, Department of Philosophy, University of California, Berkeley

Michelle Maiese, Department of Philosophy, Emmanuel College

Nico Orlandi, Department of Philosophy, University of California, Santa Cruz

Jesse Prinz, The Graduate Center, City University of New York

Anil K. Seth, Sackler Centre for Consciousness Science, University of Sussex

Lawrence A. Shapiro, Department of Philosophy, University of Wisconsin, Madison

Mog Stapleton received her PhD from the University of Edinburgh in 2012, supervised by Andy Clark. She held postdoctoral positions at the University of Stuttgart, the University of Macau, and the University of Tübingen before returning to the University of Edinburgh in 2017. Her research interests are primarily in embodied, affective, and enactive approaches to cognition.

Kim Sterelny, School of Philosophy, Australian National University

Barbara Webb, School of Informatics, University of Edinburgh

Michael Wheeler, Division of Law and Philosophy, University of Stirling

Andy Clark and His Critics

Introduction

MATTEO COLOMBO, LIZ IRVINE, AND MOG STAPLETON

Where is your mind? According to traditional wisdom in philosophy of mind and cognitive science, the machinery of your mind is just your brain. In Jerry Fodor's words: "If the mind happens in space at all, it happens somewhere north of the neck" (1999, 69).

While you are reading, your brain is buzzing with neural activity. Some patterns of neural activity support your eyes rapidly moving and then shortly fixating on the symbols in front of you. Other patterns of activity enable you to perceive, decode, and understand what you're reading. Knock down your frontal eye fields in your frontal cortex, and your eye movements will be impaired. Likewise, knock down your visual cortex, and your visual perception of what's in front of you will be seriously damaged. All this might suggest that your mind is brain-bound.

According to traditional wisdom, there is another, more fundamental reason why the physical basis of human minds cannot outrun the bounds of skin and skull. If we ask you, "What are you doing now?", you may reply along the following lines: "I want to read the introductory chapter of the volume on Andy Clark's philosophy, and I have reason to believe that moving my eyes and body in certain ways, and perceiving and deriving meaning from the symbols in front of me, is one way to get the job done." You would thus offer us an explanation of your behaviour by citing the beliefs and desires you entertain.

Importantly, the mental states cited in this explanation have a unique property. They are *about* things: they possess intentionality and have "content" (making them "semantically evaluable"). The sentences on this page have content too, but their content derives from our mental states. Sentences, and indeed any other piece of the external world, are often said to have "derived intentionality." They would not have any meaning unless it was conferred on them by people who use those sentences to express their thoughts and communicate with others. So, if external items do not mean anything on their own, and the mark of the mental is intentionality, then the mental cannot overflow the boundaries of skin and skull. And citing bits of the environment in explaining intelligent behaviour will not play the unique role played by intrinsically meaningful, content-having, causally efficacious mental states.

In sum, according to traditional wisdom, the physical substrate of the mind cannot reside outside of our heads, and explanations of intelligent behaviour must always look for content-having, causally efficacious states within the boundaries of skin and skull.

Over the course of his career, Andy Clark has systematically challenged both of these tenets in the philosophy of mind and cognitive science. Clark's challenge has primarily come on two fronts. On the metaphysical front, he has tried to show how a form of functionalism opens the door to the possibility that the vehicles of thought can overflow the causal transactions that take place within our skulls. This form of functionalism can be captured in a "Parity Principle," which Clark and Chalmers (1998, 8) formulated as follows: "If, as we confront some task, a part of the world functions as a process which, *were it done in the head*, we would have no hesitation in recognizing as part of the cognitive process, then that part of the world *is* . . . part of the cognitive process."

Here is the basic idea: if what makes something a certain type of cognitive state or process depends *not* on its internal constitution or on its physical instantiation, but rather on the coarse functional role it plays in a system, then it is a prejudice to maintain that the mind must be instantiated within the skull. When external devices, systems, and other structures in the world function sufficiently like things or processes people would normally regard as cognitive, were they to occur inside the skull, then they too could figure as proper parts of our cognitive system (Clark and Chalmers 1998; Clark 1997, 2008).

This proposal has generated a productive controversy in the metaphysics of mind, under the banner of the "Extended Mind" debate, and in the philosophy of mind and cognitive science more generally, under the banner of "Extended Cognition." In the first section of this volume David Chalmers, Fred Adams, Katalin Farkas, and Mike Wheeler articulate cutting-edge criticisms of the arguments presented in these debates, and make original proposals for the future direction of this research. Both Chalmers and Farkas focus on how exactly to specify the extended mind thesis, including how to interpret some of its early examples, and press on the possibility of consciousness extending. Adams argues that Clark's failure to specify what "cognition" is undermines his argument that cognition extends; and Wheeler pushes this style of analysis further in terms of the debates between the first and second "waves" of the extended mind debate.

Clark's arguments have also extended into a wider program investigating "Embodied, Embedded, Enactive and Affective" cognition, in which he has argued that details of the bodies of agents, as well as worldly resources, make essential contributions to explanations of a great many cases of intelligent behaviour (Clark 1997, 2004, 2008). The remaining chapters in the first section of this volume expand on this theme. Ken Aizawa pursues Clark's claims about the role of language in metacognition and related implications of extended cognition. Larry Shapiro defends the view (against Clark's "larger mechanism" framework) that the body

makes a "special contribution" to cognition. Michelle Maiese argues that affectivity is embodied and embedded, but not extended.

The hypotheses of extended, embodied, and embedded cognition have been taken up further in the informatics disciplines and in AI, in terms of Clark's proposal that humans are "natural-born cyborgs." This suggests that we are beings who routinely incorporate parts of the world into our cognition, using these not only to extend our cognitive capacities, but also to realise many novel cognitive capacities. The chapters in part 2 explore this theme.

Louise Barrett uses the theme of evolutionary (dis)continuity between humans and other animals to challenge the role of representationalism in Clark's work, and to question whether reliance on artefacts increases or decreases cognitive load. Rob Goldstone turns the Parity Principle around, and proposes that we can "hack" our own perceptual systems to provide new mental tools. David Kirsh develops a novel way to distinguish cognitive extension from cognitive embedding, and suggests that real cognitive extension is only ever brief and temporary. Kim Sterelny charts a methodologically cautious history and evolution of human reliance on materially scaffolded cognition. This concludes discussion of the first front of Clark's arguments against the received view about cognition and the nature of the mind.

On the second front, Clark has argued against classicist ideas about cognition. The classicist conception of cognition is one in which the mind functions like a digital computer, manipulating symbols according to a set of rules, where the symbols consist of concepts similar to the lexical items found in natural languages. However, Clark has argued that the functional profile of a system need not feature symbolic items that track natural language concepts, and so need not be "semantically transparent." This means that the representational items that feature in a system's processes need not relate in a systematic way to features of the world that can be picked out propositionally, with the expressive resources of public language (Clark 1989, 1993). Even systems without this semantic transparency can be justifiably taken to possess genuinely intentional states that play an important, though not exclusive, explanatory role for cognitive activities (Clark 1997, 2013).

Explanations of intelligent behaviour that rely on connectionist neural networks, dynamical system modelling, Bayesian modelling and predictive coding illuminate this explanatory shift. In these alternative approaches, contentful mental states like thoughts, beliefs and desires can be characterised as distributed sets of biases and structures in a system that is poised to pick up statistical structure in the data it encounters when it interacts with the environment. In this case, mental states need not consist of contentful symbols, but typically consist of fluid, distributed, probabilistic, and superpositional structures acquired at different time scales, which are far from semantically transparent (i.e., it's not easy to read off their meaning and how they are related to each other). And yet this opacity should not suggest that the system does not possess genuinely intentional states or internal representations, or that it cannot support the productivity and systematicity of thought.

All this may naturally suggest that we should be looking at a diverse and disunified array of explanatory resources in order to understand how minds emerge from the complex cooperation of brain processes with bodily form, action, and the canny use of environmental structures. While this conclusion seems to follow naturally from Clark's work, Clark (2013, 2016) has in fact resisted it in his most recent project. In this, he argues that the embodied mind's rag bag of tricks and stratagems may be unified through a few core principles grounded in the view that the brain is a multi-layered probabilistic prediction machine. Instead of being passive, feed-forward accumulators of environmental features, brains are active predictors of environmental signals. Hierarchically organized networks of neurons encode hierarchically organized statistical models of the environment, which they employ to make predictions about their next sensory state. As a function of the way in which observed sensory input proves these predictions to be wrong, the neurally encoded statistical models get updated and redeployed to make fresh predictions that organisms use to aptly navigate their environment. In this continuous cycle of prediction-error-based updating guided by action-perception loops, embodied brains become better and better at predicting environmental structures that matter for their own survival and flourishing.

In this exciting picture of brain function, action, perception, attention, and consciousness are painted as continuously co-constructed around the same fundamental computational routine: prediction-error minimization. Thus, our embodied, embedded, extended, and spatially and temporally distributed cognition is to be grounded in an orchestrated attempt to individually and collectively minimize the error in our predictions about specific sensorimotor trajectories in our local environments.

However, it is early days for predicting whether the predictive processing account will deliver a genuinely unified science of the embodied mind. This issue is taken up in part 3 of the volume. A number of authors in this section question the fit between predictive processing and Clark's earlier work on extended, enactive, and embodied cognition. Mike Anderson and Tony Chemero offer a different reading of how the predictive processing framework might fit with radical embodiment, which highlights the role of ecological information and downplays the concepts of prediction, models, and representation. Somewhat differently again, Karl Friston recommends an enactivist and embodied version of predictive processing in which agents need not always have (representational) models of their environments but simply *are* models of their environments in virtue of existing. Jakob Hohwy argues that the representational demands of predictive processing are not consistent with the kind of mind-world relationship that Clark is committed to. Nico Orlandi and Geoff Lee examine Clark's interpretation of the predictive processing framework and argue that it preserves traditional (and non-Clarkian) distinctions between perception and action. Jesse Prinz offers a critical overview of Clark's work and raises the question of whether it is possible for any single account of cognition to be explanatorily adequate. Anil Seth uses

the resources of predictive processing applied to interoception to explain the phenomenology of embodiment, selfhood, and subjectivity. Barbara Webb uses test cases of insect cognition to analyse whether predictive processing does in fact capture how simpler cognitive organizations such as those of insects function.

In the final section of the volume, Andy Clark offers responses to all the critics (and friends) who have contributed to this volume. His hope is to achieve, as he puts it, "a single, not wildly inconsistent, narrative."

This volume, like the work it engages with, is incredibly broad in scope and will serve to showcase and encapsulate Clark's imaginative explorations of mind and its place in nature. Andy has been a sparkly inspiration to us. Smart and curious, he has taught us how to think big and take chances, explore new intellectual territories, and face up to questions that have the potential to impact human life, and all with grace and humility. We hope that this volume will likewise inspire future generations of researchers in the sciences of mind.

> I feel released
> Bad times deceased
> My confidence has increased
> Reality is here
> The game has been disbanded
> *My mind has been expanded.*
> —Richard O'Brien, "Rose Tint My World,"
> *The Rocky Horror Picture Show* (1975)

References

Clark, A. 1989. *Microcognition: Philosophy, Cognitive Science and Parallel Distributed Processing.* MIT Press.

Clark, A. 1993. *Associative Engines: Connectionism, Concepts and Representational Change.* MIT Press.

Clark, A. 1997. *Being There: Putting Brain, Body and World Together Again.* MIT Press.

Clark, A. 2004. *Natural-Born Cyborgs: Minds, Technologies, and the Future of Human Intelligence.* Oxford University Press.

Clark, A. 2008. *Supersizing the Mind: Embodiment, Action, and Cognitive Extension.* Oxford University Press.

Clark, A. 2013. "Whatever Next? Predictive Brains, Situated Agents, and the Future of Cognitive Science." *Behavioural and Brain Science* 36:181–204.

Clark, A. 2014. *Mindware: An Introduction to the Philosophy of Cognitive Science.* 2nd ed. Oxford University Press.

Clark, A. 2016. *Surfing Uncertainty: Prediction, Action, and the Embodied Mind.* Oxford University Press.

Clark, A., and D. Chalmers. 1998. "The Extended Mind." *Analysis* 58:7–19.