BOOK REVIEW

Foundations for a coherent science of the mind

**José Luis Bermúdez: Cognitive science: An introduction
to the science of mind. Cambridge: Cambridge University Press,
2010, 576pp, £34.99 PB**

Athanassios Protopapas

Cognitive science is programmatically multidisciplinary and frequently seen as a conglomerate of disparate fields rather than a coherent approach to its purported subject matter, that is, the mind. In the singular, "cognitive science" often refers to the classical symbolic version only and not to emerging alternatives seeking to address its shortcomings. The field is vaster than the sum of its constitutive disciplines, so that becoming acquainted with cognitive science requires becoming acquainted with a multitude of theoretical and methodological approaches in addition to understanding the view of the mind as information processing and the associated body of literature in the context of this view. There are no well-worn routes to guide newcomers and no consensus as to what constitutes a core issue and what is a noncritical detail. How, then, is one to approach the task of introducing students to this field, teaching them our modes of thinking, exposing them to the main issues and findings, and ultimately enticing and attracting them as future researchers?

In short, writing an introductory textbook for cognitive science presents itself as a formidable task prior to any content or style considerations. Earlier attempts are few and far between, and present a cacophony of alternative approaches, from multiply authored, not well integrated piles of findings and opinions, to highly focused treatments of the topics best understood by the author, which fail to convey the breadth and scope of the field, presenting overly narrow views. It is very refreshing, then, to encounter a textbook that is at once broad in coverage and coherent in exposition. Bermúdez has obviously planned his approach very carefully and has made every effort to organize his selection of materials in the service of a master plan going through core issues and methods like a calendar spanning the year, including every month and week, but leaving the minutes and seconds to more appropriate devices. Intent on explanation, not mere presentation, and teaching by example,

A. Protopapas (✉)
Department of Philosophy and History of Science, University of Athens, University Campus,
15771 Athens, Greece
e-mail: aprotopapas@phs.uoa.gr

he illustrates concepts by reference to specific studies. In this way, issues and methods are covered en route to understanding theoretical approaches, providing an unforced glimpse of a multitude of research domains.

The book is organized in 14 chapters grouped into five parts. Part I, "Historical landmarks" sets the scene by going through crucial conceptual developments in psychology, linguistics, and (what is now) computer science, such as cognitive maps, phrase structure, and algorithms. The distinction among levels of presentation is given in the context of Marr's theory of vision. Bermúdez stresses multidisciplinarity from the beginning, while insisting—and ultimately demonstrating—that there is a unifying coherence to the emerging edifice. He makes sure to bring the brain into the picture early on, by introducing neural processing and intentionally conflating at this stage anatomy, connectionism, and neuroimaging, to forge the impression that multiple views of the brain are part of the route to the mind.

In part II, "the integration challenge," Bermúdez tackles integration head on, spending a section on the intertheoretical reduction that will be necessary to bring together disciplines, methods, and levels of analysis. A sensible and succinct presentation, with examples from physics (hopefully familiar to most students), helps bring home the main points—though the memory tour is comparatively less clear and of uncertain usefulness. Admirably, in the space of a few pages, we find a brief introduction to logical reasoning, modus tollens, the Wason selection task, the prisoner's dilemma, and evolutionary psychology of the cheater detection module; then a discussion of the source of the BOLD signal, instilling the notion that fMRI pictures are related to specific aspects of neuronal activity and do not indicate cognitive function in any direct or comprehensive sense. The multidisciplinarity issue is laid out as a problem of integration making up the unique character of cognitive science as a field, although the introduction of formal properties and general architectures in this section seems premature, lacking direction and closure. Bermúdez takes the opportunity to present interesting and important findings that not only exemplify integration at the two ends of the spectrum but are also worth studying on their own. What I found most appealing, in this introductory part and through the rest of the book, was the use of empirical findings in presenting, explaining, and supporting issues and concepts. Even though theories are clearly central in any presentation of cognitive science, this textbook is exceptional in not relegating experiments to the back seat. The student gains an appreciation for the foundation of cognitive science as an empirical field, building theories on a solid foundation of specific experimental findings, in their proper historical context and within the current technological reach.

Part III, "Information-processing models of the mind," deals with the two main theoretical contenders, namely physical symbol systems and connectionist networks. From the arcane problems of early artificial intelligence, the student is—somewhat abruptly—introduced to the language of thought hypothesis and propositional attitudes. Foraging on familiar turf, Bermúdez makes the problem of causation by content accessible to a wide audience and provides a lucid and remarkably even-handed presentation of the Chinese room argument and the main objections to it, leading to a clear account of symbol grounding and a less clear account of intentionality and the difference between the two. The applications of the symbolic paradigm that follow, ASCII graphics and roaming robots, seem intended

to indirectly discourage students from following the dead end of physical symbol systems by being not only simply outdated but also decidedly less interesting or attractive than apps now available on simple mobile phones. Those systems might have been interesting 30 or 40 years ago, but in 2011 they seem like black-and-white documentaries from an era long gone, both excruciatingly old and irrelevant to current cognitive research.

Connectionism is then introduced in the standard way of going through artificial neurons, single-layer networks, perceptrons, linear separability, and backpropagation, naturally leading into the biological plausibility issue. However, it seemed to me that emphasizing binary neurons and logic gates, perhaps for pedagogical purposes, might distract students from considering the importance of continuous representations and from fully appreciating the unbridgeable gap between symbolic and connectionist approaches. Connectionist applications are presented in three modeling domains: the past tense debate, naive physics (object permanence), and the balance beam problem. Each is described succinctly and is given a rule-based explanation, followed by a connectionist no-rules alternative and a specific network learning to perform the task in a way similar to children. In a great demonstration of how cognitive science proceeds, this chapter presents modeling as alternative ways of thinking about how the mind solves problems. On the other hand, it remains limited to developmental issues and may serve to reinforce the common misconception that connectionism is mainly about learning, arising from the fact that model weights are adjusted by data-driven algorithms rather than manually by design. The end of this part contains an expertly simplified exposition of the Fodor-Pylyshyn criticism, cast as an issue of "representations with separable and recombinable components."

Part IV, "the organization of the mind," begins to assemble the pieces exposed in previous chapters into more integrated presentations of theory, method, and a complete case study. The theoretical aspect begins with an overly involved discussion of the outdated Fodorian notion of modularity and intrinsic versus extrinsic properties, best left to philosophers, followed by a more appropriately leveled discussion of the massive modularity hypothesis and evolutionary psychology, without being sidetracked too much into evolution. We are then introduced into one of the best examples of successful contemporary approaches to symbolic modeling, namely ACT-R, which, however, is misleadingly presented as a "hybrid" architecture, perhaps due to a confusion about its subsymbolic level. I see nothing hybrid about distinct representational entities either occupying a slot (or a buffer) or not, as standard symbolic variables, even though certain continuous quantities may be associated with their selection. In connectionism, it is the representations per se that are inherently nondiscrete, i.e., continuous and potentially distributed.

An excellent chapter on brain mapping manages to include and discuss intelligently and coherently the main approaches to neuroimaging, from electrophysiology through EEG and MEG to PET and fMRI, carefully explaining the techniques at an appropriate level and exposing potential pitfalls to guard against in the interpretation of the results. Bermúdez does not lose the focus on integration, making clear that brain mapping is not self-serving but one potential source of information for theorizing about the mind by understanding how the brain works, appreciating the limitations of our current tools and abilities to address either goal.

The next chapter is a *tour-de-force* on theory of mind, linking pretend play and metarepresentation with mindreading, bringing together behavioral and neuroimaging data to support specific theoretical constructs and performance predictions, in an excellent application and illustration of multidisciplinary and multilevel approaches, and stressing the centrality of theory and representation. This is where students will be able to tie up various loose ends and unrefined concepts from previous chapters to appreciate the whole process at work, making theoretical progress in understanding interesting aspects of the mind.

The final part, "new horizons," is not as well thought out as the others, giving short shrift to (a not entirely accurate presentation of) dynamical systems and situated cognition approaches, which are arguably at the frontier of current developments and of potential future progress. Here, Bermúdez misses an opportunity to whet the appetite of students for exciting lines of research, dwelling on criticisms rather than opening up possibilities and presenting attractive opportunities for further thought. If a semester course makes it this far, alternative readings are to be recommended.

Overall, this is a superb introductory textbook. It would be possible, but hardly constructive, for anyone to criticize the selection of topics that were presented or left out. Bermúdez has cherry-picked theoretical topics, experimental findings, and methodological approaches, to put together a coherent and all-encompassing view of a unified science of the mind for the newcomer. In my opinion, there is nothing really crucial that is not adequately presented. The prose is admirably fluid, the content is clear, and the stated intentions of the author seem well served. I am not entirely convinced that unsuspecting freshmen or sophomores will be able to get as much out of a course based on this book as it can offer, although this would depend largely on the general level of individual institutions. I am certain that this book can open up new worlds for juniors with previous exposure to empirical sciences and developed critical thinking skills, showing that the study of the mind lies along a continuum with the rest of the physical sciences, as a distinct level of analysis on top of neuroscience qua biology, rather than in some domain of unsubstantiated verbosity and conceptual confusion coming out of armchair theorizing, as certain approaches to psychology are sometimes perceived to be. In this respect, the book does inestimable service to the future of cognitive science by exposing future researchers to clear thinking in the application of multidisciplinary methods to intriguing questions about the structure and function of the mind.

The book is also well designed from a pedagogical point of view, providing many useful devices for study and integration. Thoughtful questions are interspersed within the text, guiding students to pause and digest novel material. Boxes explain specific topics without distracting from the main flow. Annotated reading lists are provided in each chapter, wisely guiding students who wish to delve in more depth into particular issues. Helpful summaries and checklists at the end of each chapter present the main ideas briefly, in a structure that helps appreciate the flow of arguments in bare form. Most importantly, these checklists can help students confirm that they have processed the chapter successfully or indicate topics inadequately understood and in need of review.

As a first edition that has undoubtedly been in writing for a number of years, this book contains remarkably few of the usual problems, minor typographical errors,

inaccuracies, repeated material, outdated or uneven treatment, and so on, apt to be rectified in future editions. The main issue is, does the book appear well positioned to achieve the goals of instilling fundamental principles in the minds of students with no relevant background, exciting their interest and stimulating their engagement, to be further developed in more advanced courses? In my opinion, the answer is a resounding yes. This is certainly the textbook I would choose to teach an introductory course in cognitive science. Its main strengths are coherence in thought and presentation, insistence on empirical support while retaining the centrality of theory, clear exposition of all the main theoretical issues, and very helpful design to support students in studying and digesting the material. It seems that the field of cognitive science has finally been presented with a real introductory textbook that is at once comprehensive and integrated, with a clear vision of the field and a plan leading toward it.