ALVIN I. GOLDMAN

PHILOSOPHICAL APPLICATIONS OF COGNITIVE SCIENCE

FOCUS SERIES
WESTVIEW PRESS

PHILOSOPHICAL APPLICATIONS OF COGNITIVE SCIENCE

Alvin I. Goldman
UNIVERSITY OF ARIZONA

Westview Press
BOULDER • SAN FRANCISCO • OXFORD

Focus Series

Published in 1993 in the United States of America by Westview Press, Inc., 5500 Central Avenue, Boulder, Colorado 80301-2877, and in the United Kingdom by Westview Press, 36 Lonsdale Road, Summertown, Oxford OX2 7EW

Library of Congress Cataloging-in-Publication Data
Goldman, Alvin I., 1938–
Philosophical applications of cognitive science / Alvin I. Goldman
p. cm. — (Focus series)
Includes bibliographical references and index.
ISBN 0-8133-8039-1 (cloth). — ISBN 0-8133-8040-5 (pbk.)
1. Philosophy and cognitive science. I. Title. II. Series: Focus series (Westview Press)
B945.G593P55 1993
149—dc20 92-38143
CIP

Printed and bound in the United States of America

The paper used in this publication meets the requirements of the American National Standard for Permanence of Paper for Printed Library Materials Z39.48-1984.

10 9 8 7 6 5 4 3 2 1

For Raphie and Sidra

Preface

The history of philosophy is replete with dissections of the mind, its faculties, and its operations. Historical epistemologists invoked such faculties as the senses, intuition, reason, imagination, and the active and the passive intellect. They wrote of cognitive acts and processes such as judging, conceiving, abstracting, introspecting, synthesizing, and schematizing. Ethicists shared this interest in mental faculties and contents. Moral philosophers studied the appetites, the will, the passions, and the sentiments. All of these philosophers proceeded on the premise that a proper understanding of the mind is essential to many branches of philosophy. This premise is still widely accepted, but time has wrought some changes. In previous centuries the study of the mind was the private preserve of philosophy, and that is no longer true. A number of disciplines have developed a variety of scientific methods, both theoretical and experimental, for studying the mind-brain. These disciplines—the cognitive sciences—include cognitive psychology, developmental psychology, linguistics, artificial intelligence, neuroscience, and cognitive anthropology. Their practitioners attempt to understand and model the mind's wide-ranging activities, such as perception, memory, language processing, inference, choice, and motor control. Philosophy also contributes to the project, but it no longer has a privileged position.

Since it is now clear that the most detailed and reliable information about the mind will emerge from the collective efforts of the cognitive sciences, philosophy should look to those sciences for relevant information and work hand in hand with them. Cognitive science can never *replace* philosophy, since the mission of philosophy extends well beyond the description of mental processes, but it can provide a wide range of helpful fact and theory. In maintaining an alliance with cognitive science, philosophy continues its ancient quest to understand the mind. In the modern age, however, this pursuit requires careful attention to what is being learned by a new group of scientists. Plato and Aristotle created their own physics and cosmologies; contemporary metaphysicians must learn physics and cosmology from physicists. Similarly, while René Descartes and David Hume created their own theories of the mind, contemporary philosophers must give respectful attention to the findings of scientific research.

About 100 years ago, interactions between logic and philosophy assumed dramatic new importance. A similarly dramatic collaboration is now occurring between philosophy and cognitive science. When modern logic emerged in the early years of this century, philosophers saw a powerful new tool that could transform the field. Some believed that philosophy should simply *become* logical analysis, modeled, for example, on Bertrand Russell's theory of descriptions. This was doubtless an excess of zeal. But developments in logic have clearly had wide-ranging and beneficial applications throughout philosophy. Similarly, empirical studies of cognition now have great potential for enriching many areas of philosophy. This book seeks to illustrate and enlarge upon this theme.

This book does not address the methodology of cognitive science: the question of how, in detail, cognitive hypotheses or theories are tested by empirical evidence. Nor does it attempt to *survey* the various cognitive sciences. The bulk of the empirical research presented here is from cognitive psychology, but a bit is drawn from artificial intelligence, linguistics, and neuroscience. I do not attempt to give "equal time" to all of these disciplines or provide a balanced sampling of their theoretical structures. Such samplings are already available in other texts. Rather, the center of attention is the variety of philosophical problems that can benefit from cognitive studies, not the variety of cognitive studies that can contribute to philosophy. At the same time, philosophical morals are often drawn rather briefly, and the instructor or reader will often wish to pursue or debate these morals further.

My selection and discussion of material has been shaped by the desire for a short and accessible text. This constraint has dictated the exclusion of highly technical topics and topics that would require a good deal of groundwork. This is one reason there is rather scant attention to certain important areas of cognitive science, for example, the study of language. Despite such gaps, I hope that the choice of examples conveys the flavor of much of the research in cognitive science as well as its potential fruitfulness for philosophical theory and reflection.

To assist instructors and students in the further exploration of the topics covered here, I have appended a list of suggested readings at the end of each chapter. Many of these readings appear in an anthology I edited entitled *Readings in Philosophy and Cognitive Science* (MIT Press/A Bradford Book, 1993). Conceived partly as a companion to the present text, the anthology contains five chapters that closely parallel this book, plus chapters on language and methodology. I shall abbreviate citations to this anthology by [R]. When [R] follows the citation of a work in a suggested readings section, this indicates that the cited work, or some selection from it (or, occasionally, a closely related work by the same author) appears in the anthology.

I am grateful to a number of people for extremely helpful comments on the first draft of the manuscript: Paul Bloom, Owen Flanagan, Kihyeon Kim, Joseph Tolliver, and Karen Wynn as well as Westview editor Spencer Carr. Their comments resulted in numerous improvements, both substantive and stylistic.

Alvin I. Goldman

Contents

Tables and Figures

Tables

Figures

Chapter One

Epistemology

The Questions of Epistemology

Epistemology addresses such questions as: (1) What is knowledge? (2) What is rationality? and (3) What are the sources and prospects for human knowledge and rationality? To answer question 3, we would have to inquire into the specific cognitive faculties, processes, or methods that are capable of conferring knowledge or rationality. Cognitive science is clearly relevant to such an inquiry. In asking about the "prospects" for knowledge and rationality, question 3 also hints that there may be limits or failings in people's capacities to know or to be rational. Potential challenges and threats to knowledge and rationality have indeed been a focus of traditional epistemology. Here we shall address threats that stem from the potential inadequacy of some of our cognitive faculties and processes. Thus, whether we are addressing "sources" or "prospects," cognitive science, as the science of our cognitive endowments, can make important contributions to epistemology.

Knowledge and the Sources of Knowledge

Let us start with knowledge, and let us first ask what knowledge consists in. Epistemologists generally agree that knowing, at a minimum, involves having true belief. You cannot know there is a snake under the table unless you believe that there is. Further, you cannot know there is a snake under the table unless it is true, i.e., unless a snake is really there. Epistemologists also agree that mere true belief is not sufficient for knowledge, at least in any strong or robust sense of the term. Suppose you have a phobia for snakes, and you are always imagining them in this or that part of your house. You haven't looked under the table just now, nor has anybody said anything about a snake being there. But you are convinced that a snake is there. On this lone occasion you are right; someone has introduced a harmless garter snake for a practical joke. Is it correct to say that you *know* that a snake is under the table? Surely not. Thus, believing what is true is not enough to claim knowledge.

What must be added to true belief to qualify as knowledge? One popular answer, found in the *reliability* theory of knowledge, says that to be a case of knowing, a true belief must be formed by a cognitive process or method that is generally reliable, i.e., one that generally produces true beliefs. In the snake example this condition is not met. Your supposition that a snake is under the table does not stem from seeing it or from being told about it by someone who has seen it; it results from phobia-driven imagination. This way of forming beliefs is not at all reliable. Hence, although it coincidentally yields true belief on the specific occasion in question, it does not yield knowledge.

A detailed formulation of the reliability theory of knowledge requires many refinements (see Goldman 1979, 1986, 1992b). Let us suppose, however, that something along these general lines is correct. We can then return to the question posed earlier concerning the sources and prospects for human knowledge. Under the reliability theory this question becomes: Which mental faculties and procedures are capable of generating true or accurate beliefs, and which are liable to produce false or inaccurate beliefs?

In the seventeenth and eighteenth centuries, the rationalist and empiricist philosophers debated the question of which faculties were the most reliable for belief formation. The leading empiricists, John Locke, George Berkeley, and David Hume, placed primary emphasis on sense-based learning, whereas rationalists like René Descartes emphasized the superior capacity of reason to generate knowledge. Another central disagreement was over the influence of innate ideas or principles in knowledge acquisition. While the rationalists affirmed the existence of such innate factors, the empiricists denied them.

The debate between Descartes and Berkeley over the nature of depth perception will serve to illustrate this dispute. People regularly form beliefs about the relative distances of objects, but how can such judgments be accurate? What features of vision make such reliable judgment possible? As these early philosophers realized, images formed by light on the retina are essentially two-dimensional arrays. How can such two-dimensional images provide reliable cues to distance or depth?

Descartes (1637) argued that one way people ascertain the distance of objects is by means of the angles formed by straight lines running from the object seen to the eyes of the perceiver. Descartes compared this process to a blind man with a stick in each hand. When he brings the points of the sticks together at the object, he forms a triangle with one hand at each end of the base, and if he knows how far apart his hands are, and what angles the sticks make with his body, he can calculate, "by a kind of geometry *innate* in all men" (emphasis added), how far away the object is. The same geometry applies, Descartes argued, if the observer's eyes are regarded as the ends of the base of a triangle, with the straight lines that extend from them converging at the object, as shown in Figure 1.1. Thus, perceivers can compute the distances of objects by a sort of "natural geometry," knowledge of which is given innately in humankind's divinely endowed reason.

Berkeley, on the other hand, denied that geometric computations enter into the process: "I appeal to anyone's experience whether upon sight of an *object*, he computes its distance by

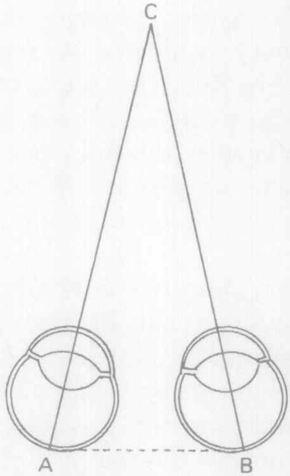

FIGURE 1.1 Schematic drawing (after Descartes) illustrating the distance information provided by convergence. Given the distance between the centers of the two retinas ($\overline{AB}$) and the eyes' angles of regard ($\angle$ CAB and $\angle$CBA), the distance of object C can be computed. *Source:* E. Spelke, "Origins of Visual Knowledge," in D. Osherson, S. Kosslyn, and J. Hollerbach, eds., *Visual Cognition and Action* (Cambridge, Mass.: MIT Press, 1990). Reprinted by permission.

the bigness of the *angle* made by the meeting of the two *optic axes?* ... In vain shall all the mathematicians in the world tell me that I perceive certain lines and angles ... so long as I myself am conscious of no such thing" (1709, sec. 12, italics in original). Berkeley held that distance (or depth) is not immediately perceived by sight but is inferred from past associations between things seen and things touched. Once these past associations are established, the visual sensations are enough to suggest the "tangible" sensations the observer would have if he were near enough to touch the object. Thus, Berkeley's empiricist account of depth perception posits learned associations rather than innate mathematical principles.

This debate about depth perception continues today in contemporary cognitive science, although several new types of cues for depth perception have been proposed. Cognitive scientists also continue to debate the role of innate factors in perception. It is widely thought that perceptual systems have some innately specified "assumptions" about the world that enable them, for the most part, to form accurate representations. An example of such an "assumption" comes from studies of visual motion perception. Wallach and O'Connell (1953) bent pieces of wire into abstract three-dimensional shapes and mounted them in succession on a rotating turntable. They placed a light behind the rotating shape so that it cast a sharp ever-changing shadow on a screen, which was observed by the subject. The shadow was a two-dimensional image varying in time. All other information was removed from sight. Looking at the shadow, however, the subjects perceived the three-dimensional form of the wire shape with no trouble at all. In fact, the perception of three-dimensional form was so strong in this situation that it was impossible for the subjects to perceive the shadow as a rubbery two-dimensional figure. From this and other studies, it has been concluded that the visual system has a built-in "rigidity assumption": Whenever a set of changing two-dimensional elements can be interpreted as a moving rigid body, the visual system interprets it that way. That is, the visual system makes the two-dimensional array *appear* as a rigid, three-dimensional body. This response can produce illusions in the laboratory, as when flashing dots on a screen are seen as a smoothly moving rigid body. Presumably, however, the world is largely populated with rigid bodies of which one catches only partial glimpses. So this rigidity assumption produces accurate visual detection *most* of the time. The assumption is innate, and it is pretty reliable.

Visual Object-Recognition

Let us further explore the prospects for vision-based knowledge by considering the way the visual system classifies objects by reference to their shape. And let us ask not simply whether such classification can be reliable, but whether it can be reliable in suboptimal or degraded circumstances, e.g., when one has only a partial glimpse of the object. After all, in everyday life things are not always in full view, and we frequently have to identify them quickly without getting a better view. Under such conditions, can vision still enable us to identify objects correctly as chairs, giraffes, or mushrooms? If so, how does it do this? Classification must ultimately proceed from retinal stimulation. But no unique pattern of retinal stimulation can be associated with a single type of object, nor even a particular instance of the type, since differences in an object's orientation can dramatically affect the retinal image. Furthermore, as just indicated, objects may be partially hidden or occluded behind other surfaces, as when viewed behind foliage. How and when can the visual system still achieve accurate object recognition?

A person stores in memory a large number of representations of various types or categories, such as *chair, giraffe, mushroom,* and so on. When perceiving an object, an observer compares its perceptual representation to the category representations, and when a "match" is found, the perceived object is judged to be an instance of that category. What needs to be explained is (1) how the categories are represented, (2) how the information from the retinal image is processed or transformed, and (3) how this processed information is compared to the stored representations so that the stimulus is assigned to the correct category.

One prominent theory, due to Irving Biederman (1987), begins with the hypothesis that each category of concrete objects is mentally represented as an arrangement of simple volumetric shapes, such as blocks, cylinders, spheres, and wedges. Each of these primitive shapes is called a *geon* (for geometrical ion). Geons can be combined by means of various relations, such as top-of, side-connected, larger-than, and so forth. Each category of objects is represented as a particular combination of related geons. For example, a cup can be represented as a cylindrical geon that is side-connected to a curved, handle-like geon, whereas a pail can be represented by the same two geons but with the handle-like geon on top of the cylindrical geon, as illustrated in Figure 1.2.

The geon theory postulates that when a viewer perceives an object, the visual system interprets the retinal stimulation in terms of geon components and their relations. If the viewer can identify only a few appropriately related geons, he may still be able to uniquely specify the stimulus if only one stored object type has that particular combination of geons. An elephant, for example, may be fully represented by nine component geons, but it may require as few as three geons in appropriate relations to be correctly identified. In other words, even a partial view of an elephant might suffice for accurate recognition if it enables the visual system to recover three geons in suitable relations.

When an object is partially occluded or its contours are somehow degraded, correct identification depends on whether the remaining contours enable the visual system to construct the right geons. Consider Figure 1.3. The left column shows five nondegraded stimulus objects. The middle column has versions of the same objects with some deleted contours. These deleted contours, however, can be reconstructed by the visual system by "filling in" smooth continuous lines. This enables the visual system to recover the relevant geons and identify the objects correctly despite the missing contours. The right column pictures versions of the same objects with different deleted segments. In these versions, the geons cannot be recovered by the visual system because the deletions omit telltale clues of the distinct geons. In the degraded cup, for example, one cannot tell that two geons are present (the bowl part and the handle). This makes identification difficult, if not impossible. Of course, one might