Boston Studies in the Philosophy of Science 282

Kathryn S. Plaisance
Thomas A.C. Reydon *Editors*

Philosophy of Behavioral Biology

Springer

Chapter 1
The Philosophy of Behavioral Biology

Kathryn S. Plaisance and Thomas A.C. Reydon

1.1 Introduction

1.1.1 Background

This volume offers a broad overview of central issues in the philosophy of behavioral biology, addressing philosophical issues that arise from the most recent scientific findings in biological research on behavior. It thus exemplifies an approach to philosophy of science that is scientifically informed as well as interdisciplinary. Accordingly, it includes chapters by professional philosophers and philosophers of science, as well as practicing scientists.

The volume originates from the conference, "Biological Explanations of Behavior: Philosophical Perspectives", held in Hannover, Germany, in June 2008. Participants in this conference represented the fields of behavioral genetics, evolutionary biology, cognitive science, philosophy of biology, philosophy of science, and communication studies. Conference presentations were organized into three main themes: explanations in behavioral genetics, developmental explanations of behavior, and the evolution of behavior. The book largely mirrors this organization, in addition to representing another theme in the philosophy of behavioral biology, namely neurobiological explanations of behavior. In what follows, we sketch out an

K.S. Plaisance (✉)
Centre for Knowledge Integration & Department of Philosophy,
200 University Ave West, University of Waterloo,
Waterloo, ON N2L 3G1, Canada
e-mail: kplaisan@uwaterloo.ca

T.A.C. Reydon
Institute of Philosophy & Center for Philosophy and Ethics of Science (ZEWW),
Leibniz Universität Hannover, Im Moore 21, 30167 Hannover, Germany
e-mail: reydon@ww.uni-hannover.de

K.S. Plaisance and T.A.C. Reydon (eds.), *Philosophy of Behavioral Biology*,
Boston Studies in the Philosophy of Science 282, DOI 10.1007/978-94-007-1951-4_1,

overview of the book, both by describing some of the major themes and philosophical context, as well as providing detailed summaries of each of the chapters.

1.1.2 Motivation & Content

One of the major motivations for this volume, and the conference that preceded it, was the observation that there were many philosophically interesting and fruitful research questions about the nature of behavior that did not fall neatly within one area, such as philosophy of biology or philosophy of psychology.[1] Thus, one must often take an interdisciplinary approach when considering scientific explanations of behavior, drawing from biology, psychology, cognitive science, anthropology, etc., and from disparate areas from within each of these disciplines. For example, with respect to psychology alone, many papers in this volume make use of and analyze research in behavioral genetics, socialization research, evolutionary psychology, and neuropsychology, to name a few. Part of our aim in this volume (and in the conference that preceded it) is to map out the philosophical domain where these different areas of work intersect and identify what might be considered the philosophy of behavioral biology.

Furthermore, just as the philosophy of behavioral biology draws on many areas of research, it also looks at a variety of behaviors in many types of organisms. With respect to humans, specific traits are considered (e.g., intelligence, personality, and schizophrenia), as well as more general behaviors such as artistic behavior or phenomena like free will and altruism. With respect to animals, scientific explanations of the development of means of communication and intentional behavior are examined. Appropriately, then, the papers in this volume reflect work from philosophers working in a wide variety of subfields – many of whom tend to take interdisciplinary approaches, and in some cases conduct their own scientific research – as well as practicing scientists (most notably evolutionary biologist David Sloan Wilson and behavioral geneticist Eric Turkheimer).[2]

1.1.3 Audience

As a result of the interdisciplinary nature of this book, we think it will be of interest to a broad audience consisting of philosophers of science, philosophers of biology, philosophers of psychology, theoretical biologists, evolutionary psychologists,

[1]Interestingly, Karola Stotz and Colin Allen make a similar point about scientific disciplines in their paper; to address this, they "promote a biologically-informed psychology and a psychologically-informed biology."

[2]In addition to including scholars from a wide range of disciplines, this volume also displays a great deal of diversity in terms of gender, nationality, and academic rank (including chapters by graduate students, postdoctoral fellows, and assistant, associate and full professors).

behavioral geneticists, cognitive scientists, and behavioral biologists more generally. In addition, some of the chapters may be of interest to those working in other areas of science or philosophy. For example, Eric Turkheimer's chapter addresses limitations of research in molecular genetics; Christine Clavien's and Rebekka Klein's discussions of altruism connect to important questions in moral philosophy; and Brian Garvey's chapter on the question of free will addresses central issues in moral philosophy as well as philosophy of mind.

For philosophers of science, including philosophers of biology and psychology, the specific papers included in this volume will be central to much of the research that focuses on philosophical issues in biological explanations of behavior, as well as more general philosophical issues such as causation and explanation. Theoretical and philosophically-minded biologists and psychologists will also find interesting and relevant work that examines the concepts, methods, and inferential reasoning of scientific research in those fields. Furthermore, some of the papers in this volume explicitly address important methodological implications of research in behavioral biology that will be of use to practicing scientists. For example, Eric Turkheimer's chapter on Genome Wide Association Studies (GWAS) could affect how research in behavioral genetics is done, and Adele Abrahamsen and Bill Bechtel's argument that the brain should be thought of as an endogenously active mechanism might lead to a new approach in neuroscience.

While the papers herein will certainly be of use to scholars in terms of their research, this volume is also intended to serve as a useful resource for teaching higher-level courses, graduate seminars, and reading groups. It offers both an overview of the issues in philosophy of behavioral biology, as well as examples of current controversies in specific fields.

1.1.4 Structure

This volume is divided into an introductory part, as well as four parts that focus on different approaches to scientific explanations of behavior: genetic, developmental, evolutionary, and neurobiological. The introductory part (Part I) includes this chapter as well as Helen Longino's paper, "Knowledge for What? Monist, Pluralist, Pragmatist Approaches to the Sciences of Behavior", which is based on her keynote address for the 2008 conference from which this book originated, and which examines and compares various approaches to the scientific study of behavior. Part II includes two papers on genetic explanations of behavior, one written by a behavioral geneticist and the other a philosopher of science, both of which focus on human behavioral traits. Part III consists of two chapters on developmental explanations of behavior, with an emphasis on non-human animal learning. Part IV is the largest section, comprising eight papers on the evolution of behavior and addressing a variety of themes such as artistic behavior, research in evolutionary psychology, and altruism. Finally, in part V, neurobiological explanations of behavior are discussed.

There are a couple of things worth noting about this structure. First, there is the obvious unevenness of the sections, with the section on evolutionary explanations of behavior dominating the various approaches, so much so that it comprises half the volume. This indicates to us a tendency in philosophy of behavioral biology to focus on evolutionary explanations over and above genetic, developmental, or neurobiological accounts.[3] Also, this particular emphasis is consistent with what we find in philosophy of biology more generally, where much of the work in that area has focused on conceptual issues in evolutionary biology.

Second, while we have been able to group the papers into these various sections in a relatively straightforward manner (by considering the approach that is the main focus of the paper), it is not the case that each paper only addresses scientific explanations that fall within that theme; rather, many of them draw on multiple approaches. For example, the paper by Karola Stotz and Colin Allen argues for a developmental approach to studying animal behavior, but in doing so they acknowledge the various roles of genetic and other kinds of factors in learning and development, arguing that these factors cannot be separated. In addition, Rebekka Klein's paper examines evolutionary explanations of altruistic behavior, though it also draws on research from neurobiology, thus providing a bridge between the two sections. As a final – and perhaps the most notable – example, Helen Longino explicitly analyzes and compares several approaches to studying behavior, including both single-factor approaches (namely those that look at the role of one type of influence, such as genetics) as well as integrative approaches that address how interactions between genetic and environmental factors influence human behavior.

1.2 Summaries of the Chapters

1.2.1 Part I: Introduction

The chapter by Helen Longino examines a variety of approaches to the scientific study of human behavior, arguing that while these approaches may at first seem to be in conflict with one another and thus amenable to comparison, each is in fact partial, focusing only on a subset of causal factors. Longino focuses on those approaches that seek to provide proximate explanations – as opposed to ultimate ones, in Ernst Mayr's terms – thus disregarding evolutionary accounts of behavior in her analysis. Instead, she focuses on single-factor approaches such as behavioral genetics, neurophysiology, and social-environmental research, as well as integrative approaches represented by developmental systems theory (DST) and gene-environment

[3] Interestingly, Helen Longino addresses a variety of scientific approaches to studying human behavior, including behavioral genetic, developmental, and neuroscientific approaches, but intentionally *excludes* evolutionary accounts in her analysis.

interaction (GxE) accounts. Given the partial nature of each approach (or their methodological difficulties, in the case of DST), Longino argues that the best way to view these different approaches is not to compare them to see which produces the correct account, but rather to take a pluralist stance. Furthermore, this pluralist stance ought to be supplemented with pragmatism, whereby one considers what kinds of questions a particular approach actually addresses, in order to make use of the knowledge that results from each of them.

Longino goes on to provide an overview of each approach, as well as some of the more important assumptions underlying them. First is behavioral genetics, which includes both classical, quantitative behavioral genetics (as seen, for example, in twin and adoption studies), as well as molecular behavioral genetics. Both of these research programs aim to identify genetic contributions to behavior, with the former estimating correlations between genetic and phenotypic variation (i.e., heritability) and the other searching for specific genes associated with the trait in question.[4] The second approach Longino discusses is neurophysiology/neuroanatomy, which aims to identify the role of neural structures and processes in behavior. This includes, for instance, studies seeking to find associations between neurotransmitters like serotonin and behavioral traits like depression and aggression. Third, Longino discusses what she calls social/environmental approaches, which aim to understand the role of environmental factors in behavior, including both macro-level variables (e.g., social class and race), as well as micro-level variables (such as family, peers, or media exposure).

Longino points out that disagreements among researchers using different approaches are not about *which* of these factors (genetic, neurophysiological, or environmental) play a role in human behavior, as they'd agree that all of them are important. Rather, they disagree about which kinds of factor are the *most* important, what methods ought to be used to estimate their relative importance, and how various interactions (e.g., between particular genotypes and environments) ought to be accounted for. As Longino puts it, "The debates, then, are less about ontology than about methodology: given that all the factors identified in the various approaches play some role, which approach is likely to be most informative about the etiology of behavior?" As she goes on to show, each approach offers something different.

The assumptions underlying each approach illustrate their partial nature. First, all assume that the behavioral traits being studied are well defined (an assumption that Longino has elsewhere critiqued at length, as she notes in her paper). Second, each of the single-factor approaches assumes that one can legitimately separate the various causes underlying human behavior.[5] Third, and most important for the

[4] Longino mentions Genome Wide Association Studies (GWAS) as an example of a popular method in molecular behavioral genetics. Interestingly, in chapter 3, Turkheimer specifically addresses recent failures of GWAS and discusses why he thinks it is not likely to be successful.

[5] This is an assumption with which Stotz and Allen disagree (see chapter 5).

purposes of taking a pluralist stance, each approach focuses on a limited range of possible causes, largely ignoring certain causal factors. As Longino puts it, "Each approach effectively situates itself in a different causal universe, making comparative assessment impossible."

One might wonder, then, whether the integrative approaches fare any better. According to Longino, while they might at first seem more promising, they encounter serious difficulties. For DST, which considers the developmental system of both the organism and its environment, separation of causes is not possible – the causal space or universe in which DST researchers work is comprehensive. As Longino points out, this is in some sense the correct picture, given the complexity of organisms (especially humans) and the development of behavioral traits. Unfortunately, with respect to human behavior at least, DST is methodologically untenable as most studies of human behavior are non-experimental for obvious practical and ethical reasons. Another, more restricted yet methodologically tractable, approach is the GxExN approach (also referred to as 'gene-environment interaction') introduced by Avshalom Caspi and Terri Moffitt. One of the main questions addressed here is, how do genes and environments interact to affect a particular neurological substrate so as to bring about a particular psychiatric disorder, such as schizophrenia or depression? For a few behavioral traits, Caspi and Moffitt have found that individuals who have *both* a particular genotype *and* have experienced a particular environmental input (such as childhood abuse) are more likely to exhibit the trait in question than those who are subject to only one type of factor. While this approach holds promise, it is also limited to disorders rather than behavioral traits subject to normal variation (such as intelligence or personality), and its findings have proven difficult to replicate.

Based on her analysis of these various approaches, Longino concludes that pluralism is the best stance to take with regard to the study of human behavior. The alternative is monism, which holds that there is only one correct account and that it is possible to figure out which one that is. However, Longino stipulates that it may be possible for many accounts to be correct, as each approach is restricted with respect to the types of explanatory factors that it can invoke, and thus one cannot legitimately compare the various accounts to one another. As she puts it, "The pluralist will propose that our task as philosophers is not to participate in debates about which of these approaches is the correct one, but to understand and help to articulate their scope, their evidential requirements, and their limitations." Longino supplements this pluralist perspective with pragmatism, arguing that the approaches discussed above should be evaluated in light of the practical goals at hand: "Pragmatism, as a second-order sorting procedure, recommends that we evaluate theories and models with respect to the specific questions they set out to answer and the kinds of intervention in the world the answers make possible." Thus, which approach we look to for answers depends on the questions we're asking and the kinds of interventions or policies that we're seeking.

1.2.2 Part II: Genetic Explanations of Behavior

The two chapters in this part of the volume examine behavioral genetic explanations of behavior, with a focus on genetic explanations of human traits. While behavioral geneticists have documented high correlations between genetic and phenotypic variance – i.e., high heritability estimates – for a number of traits like intelligence, extraversion, schizophrenia, and height, the authors point out that high heritability does not indicate how many genes are involved (Turkheimer), nor even that it makes sense to label a trait 'genetic' (Northcott). Turkheimer expands on the former point by examining the failed attempts to locate specific genes underlying heritability estimates, while Northcott draws on philosophical theories of causation to show how and why explanatory context matters in terms of whether we label a trait 'genetic'.

In chapter 3, Turkheimer notes that recent attempts to locate particular genes through Genome Wide Association Studies (GWAS) of height (a highly heritable trait) have not been very successful: only a few genetic variants have been identified, and, taken together, they only account for only about 5% of the total variation. Turkheimer explains this failure – a failure he predicted as part of his "gloomy prospect" – by demonstrating similarities between GWAS and social science and explaining why social scientists are unable to provide general causal explanations of human behavior.

Heritability estimates reflect *associations* between genetic variation and variation in a particular trait for a particular population; however, what we really want to know is whether these associations reflect an underlying *causal* process – hence the search for specific genes. As Turkheimer points out, there has been some success in identifying associations between behavioral traits (e.g., schizophrenia) and genetic variants. However, those associations have been numerous (on the order of half a million in the case of height), small (accounting for a tiny percentage of the variance), and difficult to replicate. Overall, then, they haven't added up to a causal explanation of the trait in question. More fundamentally, though, researchers have had a difficult time sorting out which of these associations are actually causal to begin with. The method used to do such sorting is null hypothesis significance testing (NHST). As Turkheimer explains, however, finding a statistically significant correlation does not guarantee a causal relationship due to the phenomenon of population stratification, where the gene variant associated with the trait in question is also associated with an environmental factor that is the actual cause (chopstick use is the classic example of population stratification). Thus, in such cases, the correlation between the genes and the behavior is a spurious one. Unfortunately, Turkheimer concludes, "NHST has not succeeded in discriminating actual causal processes from spurious correlations and non-causal associations."

Turkheimer goes on to identify an interesting analogy to GWAS in social science: the Environment Wide Association Study, or EWAS, where researchers have tried to identify the specific environmental factors underlying a behavioral trait (such as juvenile delinquency). As with GWAS, attempts to identify such specific factors have failed in the case of the environment, despite the use of a variety of statistical methods, which Turkheimer documents in detail: multiple regression (in some contexts referred to as Analysis of Covariance, or ANCOVA), Principle Component Analysis (PCA), instrumental variable regression, and propensity score analysis. As Turkheimer shows, every one of these methods is flawed in that they rely on correlations obtained from non-experimental methods, and thus cannot guarantee that any of the identified associations between a genetic or environmental factor on the one hand, and the behavioral trait on the other, indicate a causal relationship.

Despite the problems with traditional social science methods, there are some methods that are able to address the problem of population stratification, namely *quasi*-experimental designs. In behavioral genetics, these are known as within-family designs, and include twin and adoption studies. For example, by comparing the behavioral traits of pairs of monozygotic (identical) twins reared together, behavioral geneticists can obtain estimates of the nonshared environmental variance component – a measure of phenotypic differences that cannot be attributed to having different genotypes or being reared in different home environments. Interestingly, for many traits that have been studied, estimates of nonshared environmental variance have often hovered around 50%, and behavioral geneticists have tried to account for this by identifying specific environmental factors – such as differences in parental treatment, non-overlapping peer groups, unshared experiences, etc. – underlying the variation. However, like the search for specific genes using GWAS, these studies have also largely failed: many associations were identified, but none accounted for more than 2–3% of the variance component.

Turkheimer concludes, then, that GWAS is a social science as it is characterized by the following features: "1) There are a large number of potential causes, individually small in their effects. 2) The causes are non-independent and non-additive. 3) Randomized experimentation is not possible." The problem, Turkheimer goes on to explain, is not that there are many small causal factors (this is true for other areas of biology), but that the effects are interactive and thus nearly impossible to tease apart, which is made even more difficult by the fact that randomized experimentation is neither ethical nor feasible. "The problem lies in the nature of complex human behavior itself," Turkheimer observes, where the causes tend to be local and specific, rather than generalizable. By looking to social science, Turkheimer hopes that we can have a "humbler appreciation for the possibilities" of GWAS.

In chapter 4, Northcott addresses a more general issue, not of identifying specific causal factors underlying behavioral traits, but rather how we decide whether a trait is best thought of as a 'genetic trait'. Of course, as Northcott points out, every trait is a result of a complex developmental process involving a number of genetic and environmental factors. Thus, he asks, "How then can some traits usefully be termed genetic and others not?" His answer, in short, is to develop a relational definition of genetic traits that is sensitive to context; as a result, "no trait is genetic always and everywhere." Rather, whether and to what extent a trait can be counted as genetic depends on the explanatory context.

In order to develop this definition, Northcott draws on the wider causation literature from philosophy of science. In particular, he favors a contrastive theory of explanation where "a trait is genetic just in case it is explained by genes or it is *not* explained by environment. If genes made the difference, the trait is genetic; likewise, it is genetic if environment did *not* make the difference." Northcott uses the example of a trait, T, where T = his actual two legs (one of which is slightly bent due to a childhood accident). T is appropriately thought of as a genetic trait when the chosen contrast is having just one leg, but not when the contrast is having two straight legs. While this is an example of a specific (token) case, it can easily be applied to more general (type) case by considering a particular population of token cases. For example, Down's syndrome counts as a genetic trait on his definition given that there is no environmental input that could have led to its avoidance (unlike for PKU, for example). Of course, given the phenomenon of gene-environment interaction, genetic differences may only lead to particular trait differences in certain environments; thus, the same trait may or may not fall under the definition of a genetic trait depending on the explanatory context (or population of token cases) that is chosen.

As Northcott points out, one might understandably worry whether this puts too much weight on the choice of contrasts. However, as he goes on to explain, while it *is* the case that whether or not (and in what contexts) a trait is labeled 'genetic' crucially depends on the contrasts that are chosen, what matters is that the definition can be straightforwardly applied, as seen in the examples provided. Northcott concludes that, "Therefore it is not fatal that we have no foolproof algorithm for generating choice of contrast in every context. That merely implies that there may be no fact of the matter regarding whether a trait is genetic before contrasts are specified – which is exactly what a relational definition […] is claiming anyway." In other words, on this account, there just is no context-independent matter of fact as to whether a trait is genetic.

Northcott connects his account to the wider literature on causation in order to illustrate how it relates to previous work on genetic causation as well as to illuminate why labeling a trait 'genetic' might be useful. He asks, "First, consider why we should even care whether a trait is genetic or not. What normative punch could ever result form such a claim? This paper's account, by way of its connection to the causation literature, offers an answer – the counterfactuals that, according to [my definition of a genetic trait], comprise such claims are also exactly those that license *interventions*." For example, by labeling eye color 'genetic', part of what is being claimed is that no salient environmental intervention could change one's actual eye color (just as we saw above with Down's syndrome). On the other hand, while PKU was at once thought to be a genetic trait, Northcott's account suggests that it is probably *not* best described as genetic, given that there is an environmental intervention that makes a difference, namely drastically reducing the amount of phenylalanine in one's diet.

Towards the end of the chapter, Northcott introduces another important distinction: genetic traits versus genetic dispositions. As he explains, there are often contexts in which it doesn't make sense to explain the development of a particular trait *either* in terms of genes or in terms of environments – for instance, when talking about talent. He points to Mozart's musical ability as an example of a genetic disposition, as it was surely influenced by Mozart's genes but would not have been realized without his unique environment. In such cases, Northcott suggests using the terminology of a genetic *disposition* rather than a genetic trait, where the disposition "is explained (in that context) by both genes and environment but we want to focus attention just on the genes side." According to Northcott, it makes more sense to talk in terms of genetic dispositions in cases where there is a potential for an ability or talent to develop, but where that potential is only realized given a particular set of environmental inputs. He concludes by pointing out that many traits that are subject to disputes as to whether they are appropriately labeled 'genetic', such as alcoholism, schizophrenia, athletic ability, and homosexuality, are actually genetic *dispositions*, and thus that gaining a better understanding of the distinction he has introduced can help in addressing controversies over a number of human behaviors.

1.2.3 Part III: Developmental Explanations of Behavior

The two papers in this section discuss developmental approaches to the study of behavior, emphasizing learning in nonhuman animals. Karola Stotz and Colin Allen address the general conceptual relationship between learning and development, ultimately drawing conclusions not only about that relationship, but also about how behavior ought to be studied as a result. Andrew Fenton focuses his analysis on a particular organism (chimpanzees) and a particular type of behavior (evidence gathering) in order to make specific claims about nonhuman animals' status as epistemic subjects. The arguments presented in both papers have implications for scientific practice – Stotz and Allen's for comparative psychology and ethology, and Fenton's for chimpanzee cognitive studies (as well as cognitive studies of other primates).

In chapter 5, Stotz and Allen aim to examine and clarify the relationships between concepts of learning, experience, and development in the study of animal behavior. In particular, they advocate for research that integrates learning and development such that they're seen not as two separate processes (learning *and* development) but rather as part of one another (learning *as* development).

They begin by summarizing the history of the two main disciplines that study animal behavior: comparative psychology, which is situated in psychology more generally, and ethology, which stems from evolutionary biology and which has split into distinct sub-disciplines including neuroethology, behavioral ecology, cognitive ethology, and evolutionary psychology. Comparative psychologists, they point out, are largely interested in animal learning in controlled conditions, and thus favor laboratory experiments that study acquired behavior, while ethologists focus on species-typical behavior in natural habitats, often conducting field studies to examine innate behavior. Thus, this disciplinary dichotomy maps onto a dichotomy between acquired and innate, a dichotomy that Stotz and Allen reject. Despite comparative psychologists' recent claims to be taking a more integrative approach, the authors argue that they fail to take development seriously, a failure that, according to Stotz and Allen, stems from the separation of psychology from biology (with rare exceptions, such as developmental psychobiology). Thus, in order to rectify the matter, they call for a "biologically-informed psychology and a psychologically-informed biology."

Stotz and Allen go on to explicate what it means to take development seriously. Doing so, they argue, requires more than just acknowledging the importance of environmental factors; it requires the rejection of the traditional dichotomy between genetic/innate/inherited on the one hand, and environmental/learned/acquired on the other. Taking development seriously includes not just emphasizing the importance of gene-environment interaction, but refraining from drawing any distinctions between those causal factors to begin with. In support of this view, Stotz and Allen discuss various phenomena, such as developmental niche construction (where organisms actively construct their environment) and the role of environmental factors in gene expression.

The view proposed by Stotz and Allen is in line with developmental systems theory (DST), a theory they explicitly advocate and develop, in part by presenting recent scientific findings that were unavailable at the time of DST's introduction and which support and extend the theory. Furthermore, their more explicit goal is to "apply DST's framework to a new pressing question, namely how should one conceptualize the relationship between development and learning." It is just this question the authors take up in the last part of their paper by analyzing and criticizing the old distinction between learning and development.

Learning and development, they argue, are processes that should be assimilated to one another: "From a psychobiological perspective, learning appears as a category within an overall framework of development as the lifelong, adaptive construction of the phenotype out of the interaction between genes, the organism and its environment." To illustrate this, they describe various ways in which epigenetic mechanisms relate to learning and development as integrated processes.[6] Development, they say, is "the process of organismic transformation from a single cell to a differentiated, structured entity," while "learning is a specialized process of (typically neural) differentiation and structural change that supports (adaptive) modification of behavior by experience." Based on these characterizations, Stotz and Allen conclude that "learning is a kind of developmental process: i.e., learning as development" and, likewise, that development itself is a type of learning. Thus, each is a part of the other and should be studied together, rather than as distinct processes.

In chapter 6, Andrew Fenton defends the thesis that chimpanzees are "substantive epistemic subjects." In particular, he looks at two claims in support of this view: first, that chimpanzees display acts of evidence gathering, and second, that they achieve a certain amount of epistemic success in doing so. (While his analysis focuses on chimpanzees, Fenton notes that it may also be applied to other nonhuman animals.) His claim, if correct, has implications for epistemology both in

---

[6] Stotz and Allen note that most of the evidence for these mechanisms comes from experiments with animals that would not be practical or ethical to do on humans. However, they suggest that the evidence from animal studies is sufficient to warrant looking for epigenetic changes in humans, citing one study already underway that is examining the influence of parental care on child development.

that it offers an alternative to other accounts of animal knowledge (in particular, contemporary reliabilism and the anthropocentric stance), and in that it suggests the need for those working in naturalized epistemology to develop accounts of knowledge that include the epistemic activities of nonhuman animals.

Fenton relates the notion of a 'substantive epistemic subject' to Gould and Gould's account of an active knower, whereby "an organism plays an important role in the acquisition of knowledge (it learns by manipulating/experimenting with its environment)." He goes on to demonstrate how chimpanzees count as active knowers through their activities as evidence gatherers. Through a variety of detailed examples, Fenton shows the following: (1) that chimpanzees must be sensitive and responsive to changes in the social environment, and in doing so must gather evidence about social hierarchies and the like; (2) that they can acquire proto-linguistic, or perhaps even weak linguistic, skills in research settings (e.g., by learning and making use of sign language to communicate with other chimpanzees); and (3) that some of their evidence gathering behavior is analogous to that found in human children.

As Fenton explains, chimpanzee stone tool use includes moments of investigation, and more importantly, is a skill that is not innate (i.e., it is not just an expression of a genetic predisposition) but rather a *learned* behavior requiring particular environmental conditions to come about. In addition, chimpanzees' stone tool use, such as their nut-cracking behavior, requires the presence of "causally efficacious information states," which "enjoy a certain prominence in the [chimpanzee's] noetic structure." Furthermore, the knowledge (or something like it) that chimpanzees obtain from their evidence gathering can be passed on from one generation to the next – a point that Fenton suggests is relevant both to debates about chimpanzee culture and to analytic epistemologists who are interested in social knowledge.

What the detailed examples and subsequent analyses show, then, is that chimpanzees engage in epistemic activities and that these activities "track the accuracy of the relevant information states that inform the subsequent skilled behaviour." Fenton has illustrated not only that chimpanzees are evidence gatherers, but that their epistemic activities (of which evidence gathering is an example) can lead to a certain amount of epistemic success. Assuming his analysis is correct, this lends support to Fenton's original claim that chimpanzees (and possibly other nonhuman animals) are substantive epistemic subjects – a claim that has important implications for both philosophy and scientific practice.

In particular, Fenton notes that epistemologists ought to pay attention to possible cases of nonhuman animal knowledge, such as is illustrated here, and attend to the particular conceptions of knowledge, epistemic standards, and types of epistemic activities that they display. As he puts it, "If, as I have argued, chimpanzees are substantive epistemic subjects, epistemologists should not ignore their epistemic perspectives." Unfortunately, however, this is precisely what many epistemologists do. Furthermore, those who do include the epistemic perspectives of (at least some) nonhuman animals tend to treat them as *second-class* epistemic subjects, and they develop their epistemological accounts based on data drawn almost entirely from human epistemic activities. These are two practices that Fenton would like to see changed.

1.2.4 Part IV: Evolutionary Explanations of Behavior

The chapters collected in the fourth part of this volume focus on a variety of issues, including: the general question of how behavior can be accounted for from an evolutionary perspective (Wilson), evolutionary explanations of the production of art (De Smedt and De Cruz, Verpooten and Nelissen), the general research program of evolutionary psychology (Schulz, Ward), the consequences of evolutionary psychology for our conception of free will (Garvey), and evolutionary explanations of altruistic behavior (Clavien, Klein).

David Sloan Wilson opens the evolutionary part of this volume by describing some of the consequences of what he calls an "intellectual seismic shift" in thinking about human behavior. As Wilson points out, there is a long-standing tradition of thinking about human behavior as being determined by genes that are passed on faithfully from generation to generation. Genetic change takes place over a timescale of hundreds or thousands of generations, that is, on a timescale that is much larger than the duration of individual human lives and a few consecutive generations. Thus, on timescales that matter to us in our everyday lives, behavior can be considered as being fixed. However, Wilson argues, developments in evolutionary science that have been accumulating over the past two decades show that this view of human behavior as grounded in an unchanging genetic basis is mistaken.

Wilson's aim is to show how behavioral and cultural change fall within the scope of evolutionary science and to point to ways in which an evolutionary understanding of human behavior can help us to improve our lives and the societies in which we live. The phenomenon of phenotypic plasticity, i.e., the capability of organisms to change some of their traits in response to changes in the environment in which they live, occupies a central position in Wilson's argument. As Wilson points out, with respect to behavioral traits human beings exhibit a higher degree of phenotypic plasticity than do organisms of other species. This is for two reasons: First, many human behavioral traits are what Wilson calls "rigidly flexible", that is, they have a built-in capacity for providing different outputs in different circumstances. Second, many human behavioral traits can be conceived of as so-called "Darwin machines", that is, they themselves instantiate some form of evolutionary process that enables open-ended adaptation to environmental circumstances. (Although not a behavioral trait, the human immune system is a well-known example of an organismal trait that itself instantiates an evolutionary process based on variation and selection.) These two types of phenotypic plasticity, in combination with the human capability to transmit behavioral changes to later generations by means of cultural heredity and cultural evolution, render human beings highly adaptable to changes in their environments.

However, Wilson observes, the fact that evolutionary processes are a central factor in this adaptive capability of human beings, both in the form of Darwin machines within individual humans and in the form of cultural evolution, harbors both opportunities and dangers. Evolutionary processes can lead to outcomes that are beneficial to the organisms in question, as well as to outcomes that are very harmful.

Therefore, Wilson pleads that we should become "wise managers" of the evolutionary processes that concern human well-being. This management of relevant evolutionary processes, Wilson suggests, should take place by means of providing such environments in which human adaptation and cultural evolution will have the highest chance of producing outputs that further human well-being: "Provide the right conditions and the world can become a better place seemingly by itself. Provide the wrong conditions and even the most heroic efforts to make the world a better place can fail miserably." The idea is that under adverse circumstances human populations tend to evolve traits that are not conducive to human well-being: early reproduction in women and violent behavior in men, for example, are adaptations to highly insecure environments, Wilson argues. Removing such adverse conditions will lead the populations to evolve in directions that are more conducive to human well-being. What Wilson calls for, then, is the elaboration of social policies that are informed by considerations of how evolutionary processes can shape our behaviors and the societies in which we live, using evolutionary thinking for the benefit of humanity. Social policies should be aimed at providing environments that allow human populations to evolve desirable traits.

In chapters 8 and 9, a particular kind of human behavior is examined, namely the production of art. Johan De Smedt and Helen De Cruz, in chapter 8, explore the opposition between two kinds of evolutionary explanations of artistic behavior in humans: explanations that understand artistic behavior as an adaptation and explanations that see it as a byproduct of adaptations that evolved for different functions. De Smedt and De Cruz examine the evidence in favor of and the difficulties that arise with respect to both kinds of explanations and argue that in each case the problems are too large to accept the explanations in question.

In the case of adaptationist explanations, at least three problems occur. First, adaptationist explanations of artistic behavior seem all too easy to find and thus are faced with the question whether they are more than "just so" stories. Second, often such explanations are not focused precisely on artistic behavior, but attempt to explain a much broader range of behaviors, including rituals, imagination, humor, etc. That is, they don't explain artistic behavior as an adaptation, but as one aspect of a much more encompassing adaptive behavioral trait. Third, if artistic behavior is an adaptation with its own selective history, it would have to be rooted in a separate mental "art module". The modular organization of the mind, however, still is a highly problematic issue and it remains unclear to what extent the mind can actually be divided up into independently evolved modules. In this context De Smedt and De Cruz discuss results of recent neurobiological studies that raise doubt about the existence of a separate mental "art module".

Byproduct explanations of artistic behavior, however, fare no better than adaptationist explanations. According to one theory, for example, works of art appeal to human aesthetic and emotional preferences that evolved in relation to other functions. A problem with this theory, however, is that many artworks in fact don't seem to do this: De Smedt and De Cruz mention the sometimes haunting paintings by Francis Bacon as an example. In addition, the production of artworks costs considerable time and energy on behalf of the makers, the investment of which would make sense only if artistic behavior would serve a clear function (and would be an adaptation) but doesn't seem to make sense for a mere byproduct.

As an alternative to evolutionary explanations of artistic behavior as an adaptation or as a byproduct of other adaptations, De Smedt and De Cruz propose an account of artistic behavior as a product of cultural group selection. Within this framework, they explore two theoretical options: artistic behavior as a marker for altruistically/cooperatively inclined members of society (so-called "green beards") or as a marker for ethnicity, that is, for adherence to a particular set of sociocultural norms. De Smedt and De Cruz examine archaeological evidence in order to make a case for the latter option, but emphasize that probably "no silver bullet theory will be able to successfully explain all forms of art production."

In chapter 9, Jan Verpooten and Mark Nelissen present an account of the evolutionary origins of art that opposes the view presented by De Smedt and De Cruz. They address how artistic behavior can be evolutionarily explained and draw attention in this context to the importance of a particular model from sexual selection theory about the selection of signals between potential mates. They review two categories of models in sexual selection theory that can be applied to the evolution of artistic behavior, namely indirect benefit models and sensory exploitation models. According to indirect benefit models, females select males with particular traits that indicate the presence of beneficial traits in the males that they can pass on to their (and the selecting females') offspring. Such selection practices by females are indirectly selected, as they hitchhike on the direct selection of these beneficial genes. According to sensory exploitation models, particular traits may evolve if they appeal to sensory preferences of organisms that are actually aimed at different phenomena. As an example, Verpooten and Nelissen mention the evolution of orange spots in guppies: female preferences for orange food items lead males exhibiting orange spots to be more attractive to these females and thus to higher reproductive success for males with orange spots.

Although indirect benefit selection and sensory exploitation selection are usually seen as intertwined, Verpooten and Nelissen argue that at least some of the sensory biases that are found in nature might be the products of sensory exploitation selection alone. They criticize evolutionary explanations of artistic behavior framed in terms of indirect benefit models (an account proposed by Miller and one proposed by Boyd and Richerson) for underestimating the role of sensory exploitation in the evolution of artistic behavior. On the account that Verpooten and Nelissen propose, sensory exploitation is a central factor in the evolution of artistic behavior. Given that on sensory exploitation models traits evolve by means of exploiting sensory biases that evolved for different purposes, on such an account art must be understood as a spandrel, that is, a byproduct of other evolved traits – a view that clearly contrasts with the view that is defended in the preceding chapter.

Whereas chapters 8 and 9 focus on the evolution of art, chapters 10, 11, and 12 examine the research program of evolutionary psychology. In Chapter 10, Armin Schulz examines a particular strategy that evolutionary psychologists use to legitimate their approach to studying the human mind. Evolutionary psychology, especially the "strong" research program as propagated by Leda Cosmides, John Tooby, David

Buss, and others (sometimes called ‘Evolutionary Psychology’ with a capital ‘E’ and ‘P’), is regularly criticized for being too speculative in nature to be able to provide a useful contribution to the science of psychology. In particular, critics often point out that explanations in evolutionary psychology often lack a sufficient evidential basis, such that the research program rests too much on “just so” stories when trying to account for particular mental phenomena. Evolutionary psychology, critics argue, fails to provide good explanations of the sort that evolutionary biology does. One strategy of evolutionary psychologists to defend their approach is to claim that evolutionary psychology does not use evolutionary theory as a basis for explaining mental phenomena, but rather uses evolutionary theory as a heuristic tool. This defense, if adequate to the actual situation in evolutionary psychological research, would defuse the arguments of those critics who see evolutionary psychology as crucially resting on unscientific “just so” stories.

Schulz observes that this response of evolutionary psychologists to their critics is insufficiently supported, as no cases have been presented so far that unequivocally show that evolutionary theory serves as a heuristic tool in evolutionary psychology research. Schulz thus undertakes to examine the feasibility of this response, looking at evolutionary psychology in general (rather than just focusing on the “strong” program mentioned above). Interestingly, Schulz reaches diverging conclusions: On the one hand, taking Cosmides and Tooby’s explanation of cheater detection as an example, it turns out that standard examples of evolutionary psychology research do not in fact use evolutionary theory as a heuristic tool. Thus, there is strong evidence that the evolutionary psychologists’ defense fails and the critics of evolutionary psychology are right. On the other hand, however, Schulz shows that cases in which evolutionary theory plays a heuristic role can be found within evolutionary psychology, although such cases are comparatively rare. Schulz presents one such case, namely Gergely Csibra and György Gergely’s work on natural pedagogy. This case shows that there are heuristic usages of evolutionary theory in evolutionary psychology, Schulz argues, although explanatory usages – i.e., those that are subject to severe criticism – are much more common.

In Chapter 11, Chuck Ward aims to deepen a widespread criticism of the “strong” program of evolutionary psychology, namely that one of the program’s core assumptions – the assumption that the basic features of the human mind constitute adaptations to the Pleistocene environment in which our ancestors lived – cannot be upheld. Ward explores development-based criticisms of the “strong” program that focus on the phenomenon of neural plasticity, i.e., the phenomenon that the neural structures of organisms’ brains can change in response to the environments in which they live, their experiences and their actual behavior. Ward reviews evidence for the existence of neural plastic responses of the human brain to environmental cues and argues that this evidence suggests a way of explaining human cognitive processes that constitutes an alternative to the explanation that understands these processes as adaptations to life in a Stone Age environment.

Ward considers how various authors have used the phenomenon of neural plasticity as a general argument against the “strong” program in evolutionary psychology and examines two kinds of human behavior more closely: reading and writing, and musical training. These kinds of behavior constitute paradigmatic examples of culturally-mediated behaviors: behaviors that are inherited between generations because individual human beings are embedded in a common cultural environment in which they grow up and in which their neural structures develop. In recent empirical research, evidence has accumulated that practices of learning to read, to write, or to play musical instruments induce physical changes in the brains of humans involved in such practices. As Ward argues, "these examples demonstrate the existence of processes that can serve to introduce and reliably propagate modifications in our cognitive architecture *without genetic change*."

The existence of such processes, then, is inconsistent with the core assumption of the "strong" program in evolutionary psychology that contemporary human cognitive architecture has originated in the Pleistocene in the form of genetically-based adaptations and has been propagated to present-day humans by means of genetic inheritance. Contrary to the claim of proponents of the "strong" program in evolutionary psychology, Ward concludes, this program's way of explaining human cognitive traits is not the only game in town.

Brian Garvey also examines the "strong" program in evolutionary psychology, in chapter 12, but does so in relation to the issue of free will. Garvey argues that the modularity of mind that the "strong" program in evolutionary psychology assumes constitutes an obstacle to free will. According to Garvey, this obstacle is comparable to such restrictions on free will as addictions, compulsive behaviors, etc. – that is, factors that compel people to act in particular ways even if it is in principle possible for them to act differently.

Garvey discusses a number of accusations of sociobiology and the "strong" program in evolutionary psychology that interpret these research programs as implying that, if they are correct, our will is less free than we think it is. The standard defense against such accusations is compatibilism: the position that even if human actions are determined – in this case, by the makeup of our brains – it does not imply that they are not free. That is, research programs such as sociobiology or "strong" evolutionary psychology that pursue the reduction of mental phenomena to biological or physiological phenomena can be right while still leaving the possibility of having free will. But, Garvey argues, even if this compatibilist answer to the aforementioned accusations is accepted, it still may be the case that these programs provide other reasons for thinking that the human will is less free than we would think or hope.

In the case of the "strong" program in evolutionary psychology the culprit is the program's massive modularity thesis. According to this thesis, the human brain is made up of hundreds or thousands of modules, each of which has evolved in response to its own selection pressures, that is, each of which has evolved as a solution to a particular environmental problem. Moreover, proponents of the "strong" program in evolutionary psychology hold that the relevant evolutionary events occurred in the Stone Age, as the human brain has not undergone much further evolution since that time. Thus, the modules in the human brain constitute adaptations to the Stone Age environments in which our ancestors lived. Due to their being adaptations to particular environments, our brain modules make us act in ways that fit these environments. But, as Garvey notes, "what was adaptive in the Stone Age need not be adaptive now, and nor need it coincide with what we want now." Thus, if evolutionary psychologists who endorse the "strong" program are right, the makeup of the human brain causes us to have less free will than we thought we did: we are tuned to act in particular ways that might have been suitable to a different environment, but which can be thought of as a kind of compulsive behavior in our present one. In the end, however, Garvey places the burden on the proponents of approaches like the "strong" program in evolutionary psychology. As Garvey points out, while proponents of such approaches often claim that humans have free will after all, as they can override those desires that have been inherited from our Stone Age ancestors, they fail to give an account of how such compulsions might be overridden. There might thus be a way out for proponents of such approaches, but only if they provide us with the required account. As long as this issue has not been resolved, Garvey concludes, we can legitimately suspect that the "strong" program in evolutionary psychology has negative implications for our notions of free will.

In chapters 13 and 14, Christine Clavien and Rebekka Klein both address the question of how altruistic behavior can be explained against the background of human motivations that are often directed toward one's own interests and self-satisfaction. Clavien examines the opposition between two deeply entrenched positions in psychology and in the philosophy of psychology with respect to the question of whether humans are capable of behaving altruistically. On the one hand there is the position of psychological egoism, that is, the claim that human beings never perform genuinely altruistic actions; although human actions may appear to be altruistic and may have positive effects for others, humans ultimately always act in ways that are directed at their own interests. On the other hand there is psychological altruism, which holds that human beings are capable of performing genuinely altruistic actions; that is, while not all apparently altruistic actions are indeed selfless, at least some can be conceived of as being genuinely altruistic.

Clavien analyzes the long-standing debate between advocates of these opposing positions and argues that so far the debate has been carried out in an unfruitful manner: the way the debate is usually framed leads to a deadlock between the two positions, she argues, in which it is not possible to decide in favor of either of the two competitors. The central notions in Clavien's analysis of the debate are the notions of 'motive' and 'motivation'. As Clavien points out, psychological altruism and egoism are claims about the motives (i.e., psychological states, such as desires, intentions and judgments) of people that underlie their actions of helping other people. But such psychological states are extremely difficult to access and both sides in the debate can always take recourse to unconscious motives, Clavien observes. In other words, defenders of psychological egoism can always argue that, even if test persons report having motives only aimed at the interests of others, and experiments do not reveal any egoistic motives underlying apparently altruistic actions, what *ultimately* underlies the actions under consideration are unconscious egoistic motives that just fail to come to the surface. Defenders of psychological altruism can take recourse to a similar line of argumentation. Thus, "the debate over altruism cancels itself out in a battle of a priori statements," that is, a priori assumptions about empirically non-accessible, unconscious psychological states.

The way out of this deadlock, Clavien suggests, is to frame the debate in terms of the relational notion of ‘motivation’ rather than the notion of ‘motive’. A motivation is an affective state that causes someone to act; it may be based on a motive, but also on an emotion, a sensation, etc. As affective states, motivations are empirically accessible and, therefore, probably better suited to break the deadlock than are motives, which, after all, aren’t necessarily empirically accessible. Framing the debate in this way enables us to examine the role of altruistic emotions in causing apparently altruistic actions. As a consequence, evolutionary arguments can enter the debate, which, Clavien argues, is then decided in favor of psychological altruism.

Chapter 14 focuses on a particular aspect of altruistic behavior, namely punishment. In this chapter, which constitutes a bridge between Parts IV and V of this volume, Klein connects evolutionary and neurobiological explanations of altruistic and cooperative behavior with the question of how punishment may be evaluated from a moral point of view. In contrast to the psychological notion of altruism (which features in Clavien’s chapter) and the biological notion of altruism (which measures the altruistic content of an animal’s action in terms of its effect on the Darwinian fitnesses of the animal itself and of other animals), Klein focuses on the economic notion of altruism (measured in terms of the costs and benefits for the acting individual that are entailed by the action). Klein reviews results from behavioral experiments in experimental economics and research into the evolution of social cooperation that shed light on norm-enforcing practices such as altruistic punishment (i.e., acts of punishment which entail benefits for future partners in social interactions but not for the individual who performs the act of punishment, and thus function to police social interactions).

While evolutionary explanations of altruistic punishment explain why such behavior has become widespread throughout the human population, they do not explain how an individual’s motives and motivations may cause such behavior. In order to clarify this matter, Klein reviews neurobiological studies of altruistic punishment behavior which suggest that such behavior is driven by hedonic motivation and thus is connected to natural selection for the avoidance of pain and other unpleasant states. It now looks like we may have good neurobiological and evolutionary explanations of why people exhibit altruistic punishing behavior that explains such behavior as being rooted in personal motivations that are subject to selection and effects that benefit social cohesion in the group in which the individual lives (although further research is needed here). However, Klein argues, it would be too quick to value altruistic punishment as generally beneficial to society on this basis, because the personal motivations underlying punishing acts may be aimed at the welfare of society (e.g., satisfaction with keeping up the norms of society) but may also be aimed elsewhere (e.g., mere desire for revenge). Thus, Klein points out, punishment cannot generally be judged as a morally good behavior, even if the neurobiological and evolutionary explanations suggest that it furthers the welfare of society. Rather, individual acts of punishment need to be assessed by themselves, while taking into account the personal motivations underlying the act in question.

1.2.5 Part V: Neurobiological Explanations of Behavior

The two chapters collected in the final part of this volume address the question of how organismal behavior can be explained from the perspective of neurobiology. While the first of these chapters examines how neurobiologists individuate traits in need of explanation, the second chapter examines different ways of investigating brain activity and different modes of explanation associated with them.

In Chapter 15, Marcel Weber addresses two important philosophical problems that arise in relation to behavioral biology. The first is the general problem from philosophy of science regarding the theory-ladenness of observation. The second is a particular problem from the philosophy of biology, namely the problem of how organismal traits are to be individuated. In behavioral biology, Weber argues, both problems arise in connection with the identification of the explananda of behavioral biology. Behavioral biology, one might say, aims at explaining behavioral traits that organisms exhibit. But what, exactly, is a behavioral trait? What elements does a particular behavior consist of; what should be counted as part of the explanadum and what as not being a part of it; and, in particular, what determines what kind of behavioral trait a trait in question is? Organismal traits aren't simply given, but biologists have to individuate and classify them before being able to study and explain them. For behavioral traits, this is especially difficult, as behaviors often involve different parts of the organisms and often are spread out over longer periods of time.

Weber considers three ways of individuating and classifying behavioral traits: the intentional stance (according to which behavioral traits can be individuated by ascribing intentions to the animals exhibiting these traits), using proper functions to individuate traits (according to which behavioral traits can be individuated as traits that perform particular causal roles for the organisms exhibiting them, where these causal roles are the causes of the traits' presence), and using the notion of homology to individuate traits (where behavioral traits can be individuated on the basis of shared ancestry). Weber fleshes out these ways of individuating behavioral traits in more detail and ends up with five distinct theoretical notions that might constitute the basis for trait individuation and classification, but concludes that none of these does its job sufficiently well. Thus, an alternative account is needed.

Weber develops this alternative account by focusing on the notion of biological function, which in this case is conceived of by means of a version of the causal role account of functions. According to Weber, behavioral biologists (as well as biologists in other fields of work, such as experimental biology) individuate and classify the organismal traits they study on the basis of the functions that these traits perform. Here, the function of a trait is conceived of as being what a trait does or what it is capable of doing (its capacity) in the context of an encompassing system of which that trait is a part. More specifically, a trait's function is its contribution to realizing the function of the system of which it is a part. That system's function, in turn, is to be analyzed in the context of a larger encompassing system, until we reach the level of the organism. In the end, all functions are analyzed in the context of the self-reproduction of the whole organism: a trait's function (on the basis of which the trait is individuated and classified) is what it does (or is capable of doing) to contribute indirectly or directly to realizing the self-reproduction of the organism that exhibits these traits. This, Weber holds, is what makes a function into a *biological* function.

Weber supports his case by examining how behavioral traits are individuated in the study of the nematode *Caenorhabditis elegans*. *C. elegans* worms exhibit social feeding behavior which can be explained neurobiologically by means of the effect of a particular neurotransmitter in the context of the operation of a regulatory mechanism that responds to environmental stimuli. But, Weber argues, the explanandum that is being tackled neurobiologically was identified as a kind of *behavior* in the first place (instead of, for example, an instance of simple surface adhesion) by referring to the trait's function in the light of the organisms' self-reproduction. That is, while neurobiologists were examining a particular phenomenon (the clumping of *C. elegans* worms under particular conditions), it wasn't clear from the outset that this was a behavioral phenomenon – this became clear only upon consideration of the phenomenon in terms of biological functions. This way of individuating and classifying behavioral traits, however, as Weber points out, allows explananda to be changeable: neurobiologists don't just pick out phenomena in need of explanation and go on to explain them, but modify the explananda along the way.

In the volume's final chapter, Adele Abrahamsen and William Bechtel consider two different perspectives on brain and neural system activity, the reactive perspective and the endogenous perspective, and connect two modes of explanation with them. The reactive perspective on brain and neural system activity focuses on how neuronal systems respond to external stimuli. Endogenous brain and neural system activity, in contrast, is activity in absence of stimuli from the outside. As Abrahamsen and Bechtel point out, psychological and neuroscientific research has for the most part taken the former perspective, presenting test persons with specific stimuli and investigating the activity of the brain and neural system that resulted in response to these stimuli, and disregarding the endogenous activity of the brain and the neural system.

Abrahamsen and Bechtel argue, however, that both perspectives have a long history in neuroscientific research, tracing back to the late nineteenth / early twentieth century. They provide a rich historical overview of empirical research that has been done under the two perspectives, showing how the endogenous perspective has become increasingly prominent in recent neuroscientific work. In particular, Abrahamsen and Bechtel discuss recent research on endogeneous brain activity, with the aim of showing that the lack of interest in the endogeneous perspective that many neuroscientists exhibit is unwarranted. As Abrahamsen and Bechtel point out, researchers taking the reactive perspective on neuroscientific research tend to downplay the significance of the results achieved under the endogeneous perspective, treating endogeneous activity of the brain and neural system as noise rather than useful information. However, Abrahamsen and Bechtel argue, "clearly the time for dismissing the endogenous activity as mere noise has passed."

Abrahamsen and Bechtel frame the importance of the two perspectives on neuroscientific research in terms of two modes of explanation that they provide. Both are kinds of mechanistic explanations and thus fit well into what today is often called the New Mechanistic Philosophy. According to the New Mechanistic Philosophy, explanation in science often proceeds by specifying a mechanism that is capable of bringing the explanandum about. While proponents of the New Mechanicism endorse diverging conceptions of what exactly mechanisms are, Abrahamsen and Bechtel hold a specific view of what a mechanism is. They argue that two types of mechanistic explanation can be distinguished: basic mechanistic explanations, which explain by specifying the parts of a system, their organization, and the sequence the system goes through on its way to a final state from a particular initial state; and dynamic mechanistic explanations, which also include specifications of patterns of change in time that a system might exhibit. On Abrahamsen and Bechtel's account, which is an account of dynamic mechanistic explanations, a mechanism is "a structure performing a function in virtue of its component parts, component operations, and their organization," where "the orchestrated functioning of the mechanism, manifested in patterns of change over time in properties of its parts and operations, is responsible for one or more phenomena."

Abrahamsen and Bechtel conclude that the brain should be understood as an endogeneously active mechanism that is perturbed by stimuli, i.e., a system that changes its activity due to both its internal dynamics and its external perturbations; they end their chapter by arguing that the conception of mechanisms that they advocate best fits the specificities of this view of the brain. Neurobiological explanations of behavioral phenomena, then, are constructed by specifying the parts of a particular neural structure responsible for bringing about the explanandum, the properties and operations of these parts, as well as how changes in these relate to the explanandum under consideration. In such explanations, Abrahamsen and Bechtel argue, both reactions of the neural structure to external stimuli and the internal dynamics of the neural structure should be taken into account.

Acknowledgements We want to thank the 19 contributors to this volume, as well as the nearly 40 referees who generously spent their time and energy evaluating and commenting on the chapters that were submitted for publication. Also, we are indebted to those institutions who helped make possible the conference that led to this volume: the German Research Council (DFG), the Lower Saxony Ministry of Science and Culture, the Leibniz Universität Hannover Fund for Internationalization, the Center for Philosophy and Ethics of Science (ZEWW) at the Leibniz Universität Hannover, and the Minnesota Center for Philosophy of Science (MCPS) at the University of Minnesota. In addition, we're grateful to the conference participants for providing excellent presentations and stimulating discussions, as well as to an engaging audience.