An Evaluation of Means of Inquiry Into the Biological Evolution of Consciousness

A thesis submitted in partial fulfilment

of the requirements for the degree of

Doctor of Philosophy in Psychology

in the University of Canterbury by

Juliane C. Wilcke

University of Canterbury

2010

To My Grandparents
Goswinde, Immanuel, & Wurt

There probably is no more important quest in all science than the attempt to understand those very particular events in evolution by which brains worked out that special trick that has enabled them to add to the cosmic scheme of things: color, sound, pain, pleasure, and all the other facets of [conscious] mental experience.

(Nobel laureate Roger Sperry, 1964, p. 3)

Determining the best way ("method") to get what we want is the task of methodology.

(Thomas Nickles, 1987a, p. 106)

Contents

List of Figures

List of Tables

Acknowledgements

Throughout this project I have benefited from the assistance of several people. First of all, I am grateful to my supervisors, Brian Haig and Kevin Moore, for letting me go my way, while being open to meet whenever requested. Thank you for your patience when I was pursuing other academic projects and when writing was slow, for looking forward to reading each chapter, and for your comments on them. I am also thankful to Brian for his organisational and other support.

Thanks to everyone in the department who has shared the journey with me and made my research experience a bit more social. Verena needs to be mentioned here, not least because she retrieved much-needed papers from my earthquake-stricken office not long before submission. In addition, thanks to colleagues, friends, and acquaintances who have offered—or requested—to read my thesis.

I would also like to acknowledge the financial support I received for this project in the form of a University of Canterbury Doctoral Scholarship. Conference presentations of preliminary results were made possible through departmental travel funding, a New Zealand Postgraduate Study Abroad Award, and a Claude McCarthy Fellowship from the New Zealand Vice-Chancellors' Committee.

Finally, I very much appreciate the understanding that all my friends and family, both near and far, have shown as well as their willingness to carry the consequences of my pursuing this project. Special thanks go to Sabine, my mother, and Stewart, my partner, who knew what it's like and let me go for it anyway. Thank you so much for your love and for supporting my goals, regardless of the patience and sacrifices required. Stewart merits extra kudos for going through this with me day in, day out—thanks for everything!

Abstract

How can the biological evolution and functions of consciousness be studied? The purpose of this thesis was to determine not only what means of inquiry are available to do so but also how good they are or, more specifically, how promising they are with respect to the research goal of giving a scientifically respectable evolutionary explanation of consciousness. Because no suitable or easily adaptable evaluation system or set of evaluative criteria was available, I constructed a systematic tool for evaluating the promise of means of inquiry. The evaluation tool has three dimensions—relevance, efficacy, and practicality—with two criteria each, which are assessed independently (except for the relevance criteria) and synthesised into dimensional and promise scores. This tool served to evaluate, and advise on, 23 means of inquiry that have been used in the investigation of the evolution of consciousness, including its adaptation status and evolutionary functions.

The core of the thesis is formed by the evaluation tool and its application. After establishing the need for an evaluation of means of inquiry in this area and presenting the evaluation tool constructed for this purpose, I apply the tool to arguments that consciousness is an evolutionary adaptation, to general reasoning strategies, and to evolutionary strategies. This thesis core is preceded by a contextual introduction to consciousness and evolutionary theory and by the dismissal of some sceptical positions. It is followed by a comparative review of the evaluation results and an evaluation of the evaluation tool. The main contributions of this research consist of the promise evaluation tool for means of inquiry, which is underpinned by a new evaluative theory and available for use by other researchers; and, through the tool's application, an improved understanding of means of inquiry and recommendations about which of them to use for the present research goal.

Introduction

Consciousness. I'm conscious.
We [humans] are conscious.
Why? Why am I, are we conscious?

(J. C. Wilcke, personal communication, before February 20, 2003)

Consciousness is important. It is important, for instance, to each of us and to psychology as a discipline. The importance of consciousness leads to the questions of how and why it came to be in the first place. These questions address the biological evolution and potential functions of consciousness, which we can only investigate if we have appropriate means of inquiry. It is not clear that this is currently the case, as even the more developed hypotheses about the evolution and functions of consciousness are frequently based on insufficient evidence. To help advance research in the area, this doctoral thesis evaluates means of inquiry and recommends which ones are most promising for increasing our knowledge about the biological evolution of consciousness. Before giving an overview of the structure of the thesis, I explain in more detail why the topic is significant, what problem the research project addresses, and how the thesis contributes to its solution.

Brief rationale for this thesis. Consciousness is important to us individually because it is our experience of the world and of ourselves in it. Consciousness is essential to who we are. Without consciousness humans are not only unaware of anything, but also inactive: People who are in a dreamless sleep or coma normally neither obtain nor consume food and drink; they do not protect themselves, interact socially, or have sex. And although many bodily functions are unaffected by the absence of consciousness, prolonged unconsciousness without intervention leads to death within a few weeks at most. This is not to say that consciousness is a prerequisite for human existence in principle; there may well be ways to make us survive and reproduce without consciousness. It is just that *we* cannot live our lives naturally without consciousness.

The personal centrality of consciousness makes the topic very relevant, yet its study difficult.

Consciousness is also important to psychology. Over 130 years ago, when psychology became established as an independent empirical discipline, psychology *was* the scientific study of conscious experience. Although the scope of psychology has expanded and now includes all mind and behaviour, consciousness has been repeatedly deemed its defining domain. The influential psychologist George Miller, for instance, explained in an interview that he took consciousness to be "the constitutive problem of psychology" (Baars, 1986, p. 220). One reason for this view is that an understanding of consciousness is needed to understand unconscious mental processes, either or both of which are involved in all mental phenomena and therefore essential to their explanation (Searle, 1998). Another reason is that consciousness could help to integrate psychological theories from diverse research areas (Banks & Farber, 2003): It makes a difference in perception, learning, memory, thinking, attention, emotion, volition, motor control, and attitudes; and corresponding mental contents appear unified in consciousness. Despite the varied treatment of consciousness as a research topic in its own right, it has been an "ever-present concern for everyone thinking about the human mind" (Güzeldere, 1995a, p. 36). This is not surprising because we tend to consciously experience the stream-of-consciousness I as that part of us that is perceiving, thinking, feeling, and acting, to the exclusion of unconscious processes. Psychology cannot exist without consciousness (Marcel, 1988; Roback, 1952), if only because much of psychology is performed by conscious psychologists and relies on conscious reporting of conscious contents, both in research and practice. Consciousness is thus not the only topic in modern psychology, but it matters fundamentally to several aspects and research areas of the science of mind and behaviour.

Given the importance of consciousness to individual human beings and its significance within psychology, it seems natural to ask why consciousness exists. Assuming that inanimate matter on the early Earth was not conscious, consciousness must have arisen since then during chemical or, more likely, biological evolution. The questions of where consciousness has come from, and why, have become more pressing in the last few decades with the discovery of an increasing number of unconscious mental processes and behaviours previously thought to be tied to consciousness (Dennett, 1987; Frith & Rees, 2007). In a recent survey of the top 25 big questions facing science over the next quarter-century, the journal *Science* noted that, "ultimately, scientists would like to understand not just the biological basis of consciousness but also why it exists. What selection pressure led to its development, and how many of our fellow creatures share it?" (G. Miller, 2005, p. 79). An evolutionary approach to consciousness is necessary to answer such questions and may additionally be useful for deciding between philosophical theories of consciousness (Bechtel & Richardson, 1983; Carruthers, 2000). Moreover, knowledge of the relevant functions of consciousness is likely to have implications for psychotherapy, neuropsychology, medicine, ethics, and law (Haynes, Roth, Schwegler, & Stadler, 1998). One major route, then, to finding out why we have conscious experiences is to study the phylogenetic origin and evolutionary history of consciousness, including potential evolutionary functions.

What stands out about the literature on the evolution and functions of consciousness is that, by and large, it consists of many diverse hypotheses, the majority of which are insufficiently supported. This literature has been characterised as "surprisingly barren" (Güzeldere, 1995b, p. 131) with still only "little understanding about how it [consciousness] may have evolved through time" (Nielsen & Day, 1999, p. 95). It appears that authors recognise the importance of the topic and, hence, feel their accounts of consciousness would be incomplete without some comments on the topic, but find it either too obvious or too difficult, and not their main purpose anyway. Of course there are exceptions, and fortunately the situation has begun to improve: Several recent publications dedicated to the topic have used theoretical and empirical arguments to make progress on discovering how and why consciousness evolved (e.g., Cabanac, Cabanac, & Parent, 2009; Ginsburg & Jablonka, 2007a, 2007b; Merker, 2005; Morsella, 2005). More interest in the evolution and functions of consciousness, together with a growing number of researchers prepared to tackle the topic seriously, set the course for the evidence-based development of associated hypotheses.

So how can willing researchers investigate the biological evolution of consciousness? In order to add to our limited knowledge in this area and to solve some of the problems apparent in the literature, we need to know how we can obtain support for, or against, hypotheses about evolutionarily relevant functions of consciousness or other aspects of its evolution. There are three main reasons why means of inquiry deserve special attention in the present context. To begin with, some aspects of the study of consciousness are methodologically difficult (Banks & Farber, 2003; Velmans, 1993; Wright, 2007). Another reason is that it is simply not obvious in this emerging interdisciplinary research area which strategies and methods are best for studying the evolution of consciousness. An even more important reason for focusing on means of inquiry is that those means that have been employed so far differ greatly in their usefulness. Considering the increasing research interest in the evolution and functions of consciousness, it is timely to examine the available means of inquiry and to determine their value systematically. This doctoral thesis presents just that: an evaluation of means of inquiry into the biological evolution of consciousness.

Means of inquiry and methodology. Because means of inquiry are central to the present project, it is worth explaining what they are and which ones are evaluated here. *Means of inquiry* is a broad term, encompassing many different procedures and tools that are employed by researchers to attain those research objectives, at different levels of generality, that are directly related to learning more about the phenomena under investigation (i.e., excluding research objectives such as the dissemination of findings and enabling conditions, e.g., time, funding, and ethics approval). Examples of means of inquiry are general research approaches, research designs, research methods for data collection and analysis, technical instruments, and certain cognitive processes. Given the current state of the literature on the evolution and functions of consciousness, the evaluation reported in this thesis was designed to concentrate on research strategies, research methods, and arguments.

The thesis is thus in research methodology, the study of research methods or, more generally, the study of means of inquiry. It is not concerned with general accounts of scientific method, such as the hypothetico-deductive and inductive theories of method. Instead, the thesis is concerned with the evaluation of more context-specific means of inquiry themselves. The evaluation is underpinned by a broader account of scientific method than the aforementioned ones, namely the abductive theory of scientific method (Haig, 2005). One of the major strengths of this theory is that it accommodates means of inquiry used in both the detection of empirical phenomena and the construction of explanatory theories, all means of which contribute to scientific progress.

Within the present context the thesis performs the three major tasks of methodology (Nickles, 1987b, 1989): It describes, criticises, and advises on means of inquiry into the evolution of consciousness, with an emphasis on their evaluation. Methodology can inform crucial research decisions in this way, such as the choice of particular means of inquiry for a given research problem. The means of inquiry that we use not only determine to what extent a chosen research goal may be reached and enable us to do so; they also determine to what extent the resulting claims about empirical phenomena are justified on the basis of having been acquired by reliable processes (Goldman, 1979; Haig, 2005). Textbooks typically instruct their readers to match the research method to the research question. Yet additional guidance is required because different methods can result in different answers to the same research question, a problem known to psychologists as *method variance* (D. T. Campbell & Fiske, 1959), and because there are many other factors that influence method choices (Golden, 1976; Kulka, 1981). This thesis provides new systematic information on means of inquiry in the present context and, on that basis, recommends which of them are most promising for contributing substantially to an evolutionary explanation of consciousness.

A clarification of the scope and focus of the thesis is in order here. The thesis is not a complete survey of means of inquiry that have been, or could be, used to investigate the evolution or functions of consciousness; however, it does cover a wide range of means of inquiry as representative examples. Similarly, the thesis is not intended as an exhaustive review of suggested functions of consciousness or related evolutionary hypotheses, although it does cite many of the corresponding references. (Fortunately, readers interested in the former topic can consult two recent overviews: A. K. Seth, 2009, and Van Gulick, 2009.) This is because the focus of the thesis is on the evaluation of means of inquiry, not on the content of hypotheses about the evolution and functions of consciousness. In particular, the thesis evaluates the promise of means of inquiry with respect to the research goal of giving a *scientifically respectable evolutionary explanation of consciousness*, using a tool constructed specifically for this purpose.

The construction of an evaluation tool for means of inquiry was necessary because I could not find any evaluation system in the literature that was either suitable for the present research project or that could easily be adapted to it. Because this thesis is about the first stage of an evaluation tool's construction in a new content area (i.e., the evaluation of means of inquiry for a particular research goal), I treated this project as an exercise in evaluation rather than in psychometrics. That is, attention to the foundations of evaluation was deemed of higher priority at this stage of research than a strong focus on psychometric virtues such as test validity and reliability. Psychometric considerations will become more important once the tool has proved practical and worthy of such development. The tool's construction in the thesis was therefore guided by the logic of evaluation and corresponding methodological resources.

The main contributions of this research stem from the construction and application of the evaluation tool for means of inquiry. Both of these components, the tool and its application, contribute knowledge for instrumental and conceptual use (Cousins & Shulha, 2006; Rich, 1977), so-called knowledge for action and knowledge for understanding. The evaluation tool is based on a new evaluative theory of the promise of means of inquiry in the present context, itself intended as a significant contribution to the meagre literature on the evaluation of research methods. The tool is available for use by other researchers either here or, after adaptation of the instructions in a context-specific dimension, in other areas. Its application in this thesis leads to a better understanding of the means of inquiry in the present context, which has the potential to improve their future use and to inform their development. Last but not least, the resulting recommendations on the promise of means of inquiry, if followed, can facilitate progress in research on the evolution and functions of consciousness. In sum, this thesis contributes instrumental knowledge in the form of a method evaluation tool and as recommendations on means of inquiry into the biological evolution of consciousness; it also contributes conceptual knowledge, namely, an evaluative theory and an improved understanding of means of inquiry.

How this thesis is structured. I first give an overview of each chapter's role in the thesis and then describe individual chapters in more detail. The core of the thesis, which spans four chapters, is formed by the evaluation tool for means of inquiry and its application to research into the biological evolution of consciousness. Chapter 2 establishes the need for an evaluation of means of inquiry in this area and presents the evaluation tool constructed for this purpose. The tool is then applied to means that have been used to study the evolution and functions of consciousness, in particular: to arguments that consciousness is an evolutionary adaptation in Chapter 3, to general reasoning strategies in Chapter 4, and to evolutionary methods in Chapter 5. This thesis core is preceded by a contextual introduction to consciousness and evolutionary theory and by the dismissal of some sceptical positions in Chapter 1. The core is followed by a comparative review of the evaluation results, some suggestions from evolutionary psychology, an evaluation of the evaluation tool, and conclusions drawn from the research in Chapter 6.

Chapter 1 begins with the question of what consciousness is. Some knowledge of the subject matter is needed to avoid misunderstandings and to understand the methodological challenges faced when studying its biological evolution. It is also needed directly in the evaluation of means of inquiry, namely, for determining which aspect of the research goal a means addresses. The other ingredient of the means' research goal is evolution, so the chapter discusses evolutionary explanations next. This chapter section serves to establish common ground on what evolution is taken to involve and what evolutionary explanations should look like. In order to allow for the possibility that consciousness is functional, the following section rejects epiphenomenalism, the view that mental events do not influence physical events. The chapter also dismisses other doubts about research on the biological evolution of consciousness, including in-essentialism, cognitive closure, and culture as an alternative source of consciousness. Chapter 1 thus functions to clear the way for the method evaluation tool and its application by providing the necessary background and dealing with general doubts about the research programme.

Before Chapter 2 introduces the evaluation tool for means of inquiry, problems in the literature on the biological evolution and functions of consciousness are identified. As already mentioned, current hypotheses tend to be inadequately supported, but there are other problems too. The following chapter section explains how an evaluation of the promise of means of inquiry can help to solve some of the methodological difficulties evident in the literature. The construction of the evaluation tool is then described, including its three evaluation dimensions, the choice of evaluation criteria, and the scoring scale. The chapter also specifies the tool's rubrics, which are used to assess each means of inquiry on the criteria, and its synthesis procedure for arriving at dimension and promise scores. The description of the evaluation tool, itself an outcome of the research project, acts as the method section for the evaluation of means of inquiry.

The three results chapters report the application of the evaluation tool to means of inquiry that have been used to study the evolution or functions of consciousness. Chapter 3 deals with means employed to support the hypothesis that consciousness is an evolutionary adaptation, independent of the function it may have. Ten arguments are evaluated, some of which are based on subjective experience, others on casual evolutionary reasons; the most promising ones are from evolutionary biology. Chapter 4 is devoted to the evaluation of common general strategies in the present context, such as input from introspection and folk psychology and also more abstract reasoning strategies. *Contrastive analysis* (Baars, 1988) for determining the functions of consciousness is assessed in this chapter, as is the related strategy of excluding unconsciously performed functions. The final of the evaluation chapters, Chapter 5, turns to methods with a clearer evolutionary focus. It first evaluates the *natural method* (Flanagan, 1992), including its use for generating and developing "deflationary evolutionary explanations" (Polger, 2007, p. 83). Because most of the remaining strategies take long-term developments of consciousness into account, comparative methods, indicators from brain evolution, and other such means of inquiry promise to contribute much to an evolutionary explanation of consciousness.

The concluding chapter of the thesis, Chapter 6, begins with a comparative review of the evaluated means of inquiry and their promise. The following section considers what we can learn from the field of evolutionary psychology for better understanding and improving research on the biological evolution of consciousness. In addition to points such as the importance of alternative hypotheses and realistic expectations, it mentions means of inquiry that have not yet been used in the present context. The chapter then evaluates the evaluation tool, with a focus on difficulties that became apparent during its application and solutions to them. It ends with an overview of the major contributions of the research reported here and lists recommendations for future research in the hope of making progress toward a scientifically respectable evolutionary explanation of consciousness.

1 Approaching consciousness confidently from evolutionary biology

Even if "consciousness is the biggest mystery" (Chalmers, 1996, p. xi), we need to engage with it to find out whether we can increase our understanding of it. Before looking at how we can study consciousness, it is worth getting a better idea of what it is we are dealing with: What may researchers have in mind when they use the term *consciousness*, and are there different kinds of consciousness? Because later chapters evaluate means of inquiry into the biological evolution of consciousness, we should also determine whether it is at all reasonable to approach consciousness within an evolutionary framework: Are there any serious indications that biological evolution is not relevant to the study of consciousness or that such an approach is unlikely to succeed? Overall, this chapter introduces what we are getting ourselves into and provides the necessary background for the thesis.

The chapter begins with a discussion of the definition of *consciousness* and a description of important distinctions, before specifying the main assumptions that this thesis makes about consciousness. Note that the study of consciousness and associated methodological worries are not addressed in detail until the following chapter. In the second section of the present chapter, I give a brief overview of evolutionary theory, touch on what an evolutionary explanation of consciousness might have to involve, and clarify my view of evolutionary psychology. I then consider arguments for and against epiphenomenalism, which is in this context the philosophical theory that consciousness does not affect any physical events and, hence, cannot have a function. The last section of the chapter deals with three additional concerns: that consciousness is not necessary in principle for any of the functions it may perform, that consciousness is predominantly a social product, and that understanding consciousness is beyond our cognitive abilities.

1.1 Consciousness in this thesis

The answer to the question "What in the world can consciousness be?" (Dennett, 1987, p. 160) is often treated as either very obvious or very difficult. On the one hand, conscious experience seems "the most familiar thing in the world" (Chalmers, 1996, p. 3), which is presumably why Sigmund Freud (1933) commented that "what is meant by 'conscious', we need not discuss; it is beyond all doubt" (p. 94). Some psychology dictionaries, on the other hand, assert that a definition of *consciousness* is impossible (Sutherland, 1989), or advise psychologists to leave the term's description to philosophers (Moulin, 2006). Both of these attitudes toward *consciousness* leave us, at least for now, without an explication of the term.

We do, in fact, not need a precise definition of *consciousness* in order to increase our knowledge of it. Analytic definitions, which aim to identify the essence of a concept, tend to come late in scientific investigations, not early (P. S. Churchland, 1988; Searle, 1998; Weiskrantz, 1988). This is an old insight: Aristotle regarded such definitions as the aim of knowledge generation (Popper, 1945), and logicians in the middle ages assigned them to the final step in the progress of knowledge, a view that was largely borne out by the middle of the 19th century (Whewell, 1840). What is more, Karl Popper (1945) argued at this university that science, instead of overburdening its terms with precise meaning, concentrates on the precision of its theories and only introduces explicit definitions as shorthand labels. Accordingly, Crick and Koch (1990) dismissed attempts at formally defining *consciousness* as premature until we know much more about it.

What we do need, if we want to study the same phenomenon, is a common understanding of the term *consciousness*. I doubt that it is sufficient to state that "everyone has a rough idea of what is meant by consciousness" (Crick & Koch, 1990, p. 264) and to leave it at that. People do not agree on what consciousness is (Kemp & Strongman, 1994), and not even researchers in the same discipline may mean the same thing when they use the term *consciousness* (Chalmers, 1995; Güzeldere, 1995a). To provide a useful starting point, I quote two descriptions of consciousness. Both give synonyms, followed either by conditions of occurrence: "states of sentience or awareness that typically begin when we wake from dreamless sleep and continue through the day until we fall asleep again, die, go into a coma, or otherwise become 'unconscious' " (Searle, 1998, p. 24); or by contents of consciousness: "inner experience ... what it *feels* like to be oneself, to have sensations, thoughts, moods, desires, subjective reasons for one's actions" (Humphrey, 1987b, para. 3). These aspects may be combined by describing consciousness as "the normal mental condition of the waking state of humans, characterized by the experience of perceptions, thoughts, feelings, awareness of the external world, and often in humans (but not necessarily in other animals) self-awareness" (Colman, 2009). As a first approximation, then, consciousness can be described as an organism's awareness of the environment and probably, at least at a very basic level, of itself in it.

Distinctions within consciousness. The identification and examination of various concepts and aspects of consciousness have helped to improve our understanding of both the construct and the phenomenon itself. In line with this, Carruthers (2000) called the distinction of several notions of the term "one of the real advances made in recent years" (p. 254). Furthermore, distinguishing different kinds of consciousness facilitates phylogenetic approaches by changing the focus from human consciousness in its entirety to the question of which species have which kinds of consciousness (Rensink, 2009). I briefly describe the most relevant concepts and aspects of consciousness below. The purpose of this description is not merely to clarify the uses of terms in the literature: Which aspect of consciousness a means of inquiry serves to address is important for the means's evaluation in this thesis. For example, the addressed aspect matters when assessing the potential contribution of a means to the research ideal, that is, to explaining the biological evolution of all aspects of consciousness in all organisms. Note that some concepts of consciousness are closely associated with particular theories of consciousness, the description of which would take us too far afield without leading to a coherent framework. I therefore adopt the authors' own terms for consciousness and its aspects in the thesis and comment only on seemingly unusual usage.

A breakthrough in the study of consciousness was the realisation that we could compare conscious and unconscious mental events (Baars, 1997a, 2003b). Yet this dichotomy turns out to be a simplification too. For instance, instead of two different types of mental processing, we may usefully distinguish at least three: subliminal, preconscious, and conscious (Dehaene, Changeux, Naccache, Sackur, & Sergent, 2006), or nonconscious, conscious, and metaconscious (Schooler, 2002). Aside from mental processes, organisms can also have different levels of consciousness at different times, spanning unconscious, sleeping, drowsy, relaxed, and alert states. Moreover, consciousness could vary along one or more continua rather than consist of discrete levels. Suggested dimensions include, for example, clearness of experience (Ramsøy & Overgaard, 2004), richness of content and degree of influence (Dennett, 1995b), and strength, stability, and distinctiveness of representations (Cleeremans & Jiménez, 2002). Whether consciousness and its aspects are indeed continuous or discrete is still under investigation (e.g., Christensen, Ramsøy, Lund, Madsen, & Rowe, 2006; Dupoux, de Gardelle, & Kouider, 2008; Overgaard, Rote, Mouridsen, & Ramsøy, 2006; Sergent & Dehaene, 2004). And, as illustrated by the example of cooling water turning into ice (Cleeremans, 2005), these two possibilities do not exclude each other.

There are not only different levels of consciousness but also many different contents of consciousness. We can be conscious of perceptions, thoughts, emotions, memories, mental imagery, and so forth (more precisely, we can be conscious of their objects and of having or doing them); and each of these categories has many subcategories in turn. Taking perception as an example, there are visual and auditory experiences, tastes, smells, touches, pains, experiences of temperature and the location of body parts, and several other types of perceptions. Even these subcategories may divide further, for instance, into sweet, bitter, sour, salty, and umami tastes. Experiences can thus differ between categories, but they can also differ considerably within a single category. The sweetness of a banana, for example, is very unlike that of its foam-candy version. Another factor that contributes to the great variety of conscious contents is that a particular conscious experience normally consists of a combination of experiences from different categories. While reading these sentences you may be aware of my phrases on paper or a computer monitor and of your understanding of them, as well as of associated thoughts, emotional reactions, and distracting sounds in the background. Tononi and Edelman (1998) characterised conscious experience accordingly as both integrated and differentiated from billions of other possible conscious states.

The worth of distinguishing between these two aspects of consciousness, levels and contents, is well accepted (e.g., Laureys, 2005; A. K. Seth, 2009). When one of them is constant or otherwise not of interest, it is reasonable to focus on the other aspect. This is what Searle (1998) and Humphrey (1987b) did in their descriptions of consciousness quoted above. However, in other situations it may be more appropriate to consider both aspects. For instance, altered states of consciousness, which include mental states during dreaming, meditation, hypnosis, psychosis, coma, orgasm, and those influenced by psychoactive drugs, are likely to differ in their level and in their contents of consciousness. With these two basic aspects of consciousness explained, we now turn to three further distinctions, which reappear later in this thesis.

First, Rosenthal (1986, 1993, 2009) has argued that the following three concepts of *consciousness* need to be distinguished because they refer to distinct phenomena. The first concept is *creature consciousness*, which denotes an organism's being awake and responsive to sensory stimulation. The second concept, *transitive consciousness*, refers to an organism's being conscious of something, either by perceiving it or by thinking about it. (It is called *transitive* because the phrase *to be conscious of* takes a direct object.) These perceptions and thoughts may themselves be conscious or not, leading to the introduction of a third concept: *state consciousness*, which denotes a mental state's being conscious, as opposed to unconscious, and is the subject of much recent research and debate. An organism is conscious of such mental states—in a way that depends on the particular theory—with or without a deliberate introspective focus on them. If the phenomena corresponding to these three concepts of *consciousness* have different evolutionary functions, including none, as Rosenthal (2008) maintained, their distinction is clearly relevant to this thesis.

Second, the most widely known (and disputed) distinction is that between *phenomenal* and *access consciousness*, which are part of Block's (1995) conceptual quartet. He describes phenomenal consciousness as experience, which occurs, for example, when humans see, hear, or feel, but also when they think or want. What is specific to phenomenally conscious states is that "there is something it is like to *be*" (Nagel, 1974, p. 436) in them. A state is access conscious, in contrast, if its content is "poised for free use in reasoning or in rational control of action" (Block, 1995, p. 238). The best indicator of access consciousness is often reportability. Block's other two concepts are *self-consciousness* and *monitoring consciousness*, both of which apply to organisms rather than mental states. Animals are self-conscious if they have a self-concept and can use it in thinking about themselves, as suggested by, for instance, mirror self-recognition. Monitoring consciousness, which is also called *reflective consciousness*, may be a form of inner perception, internal scanning, or higher-order thought. These distinctions are important here, as Block's discussion of them aims to expose faulty reasoning about the functions of consciousness.

Third, primary consciousness differs from higher-order consciousness (G. M. Edelman, 1989, 2003; G. M. Edelman & Tononi, 2000). An animal with primary consciousness has the ability to generate a multimodal phenomenal scene in the present, which integrates perception, motor acts, and memory. Higher-order consciousness requires primary consciousness, as well as semantic capability and, for its most developed form, linguistic capability. It allows the explicit construction of scenes in the past and the future, and it comes with a sense of self (self-consciousness) and consciousness of being conscious (reflective consciousness). This means that only animals with higher-order consciousness can report their conscious experiences. The distinction between primary and higher-order consciousness is supported by the existence of related terms, such as *sensory awareness* and *meta-consciousness*, and similar distinctions (e.g., Damasio, 1998; Dennett, 1987; Ehrlich, 2000; Johnson-Laird, 1988; Macphail, 1998; O'Hear, 1997). What matters here is that, if primary consciousness is a prerequisite for higher-order consciousness, it probably arose earlier in biological evolution, and that, if both have evolutionary functions, we can expect these to differ.

By now it should not come as a surprise that *consciousness* is frequently considered an umbrella term (e.g., Gillett, 1988; Van Gulick, 2009), which covers "a rag-bag of sundry effects" (P. S. Churchland, 1988, p. 281) or refers to a "heterogeneous hodgepodge" (Flanagan, 1991, p. 361). Does this indicate that consciousness is not one thing but many? To begin with, it is important to remember that all phenomena subsumed under the term *consciousness* have in common that they are experienced consciously (Flanagan, 1991). Yet this shared property may reflect our particular intuitions more than any real commonality in the phenomena of interest. This possibility is suggested by the late emergence of the term *consciousness* with its present range of meanings in English, in the 17th century, and by the lack of a corresponding term in ancient Greek and in Chinese (Wilkes, 1988). It is additionally supported by the general tendency of folk theories to categorise phenomena by appearances, combined with the common recategorisation of these phenomena once a better scientific understanding of them has been achieved (P. S. Churchland, 1988). We should certainly be aware of the possibility defended by Wilkes (1984, 1988) that not all conscious phenomena are linked in a scientifically interesting way. However, I agree with P. S. Churchland (1988) and Van Gulick (2009) that we simply do not know whether *consciousness* will turn out to be a scientifically meaningful term.

Let us turn to what the heterogeneity of consciousness means for research on the biological evolution and potential functions of consciousness. As already pointed out for the three main distinctions above, each kind of conscious phenomenon is likely to have its own evolutionary history and functional role, if any (Carruthers, 2000;

Flanagan, 1995a; Polger & Flanagan, 2002; Van Gulick, 2009). Consequently, the research goal of giving an evolutionary explanation of consciousness requires us to focus on precisely those kinds of consciousness that are evolutionarily significant. This is difficult because we do not know the relevant demarcations of consciousness yet, but we can make educated guesses, specify what kind of consciousness we are addressing (Van Gulick, 2009), and expect the demarcation to become clearer as we learn more about the neural bases, functions, and evolution of consciousness (Cartmill, 2000; Frith & Rees, 2007; Polger & Flanagan, 2002). Because these research areas are strongly interrelated, we should not defer investigating the evolution of consciousness for lack of a definitive classification, but rather let progress in each research area inform the others. Determining what kinds of consciousness figured in its evolution is part of answering the research question about the evolution of consciousness.

Assumptions, limitations, and conventions. To clarify my approach in this thesis, I outline below its basic assumptions about consciousness within a scientific framework. Because these points seem sufficiently uncontroversial, at least among the majority of scientists working on consciousness, I do not argue for them in detail. In addition, I point out related topics that are not covered in the thesis and introduce some conventions.

- ***Scientific realism.*** Science aims to construct true theories of observable and unobservable phenomena. These phenomena exist in the real world, whether or not researchers study them. Science is frequently successful in advancing towards (approximately) true knowledge about the investigated phenomena.

- ***Consciousness realism.*** When we speak of consciousness, we are referring to an existing phenomenon; consciousness is not merely a belief, a way of speaking, a social construction, or a theoretical tool without a real-world referent. While our knowledge about consciousness is limited and tentative, consciousness realism is a working assumption. Modifications of the concept and its subcategories are to be expected with scientific progress. For example, some seemingly important features of consciousness, such as the stream of consciousness (Blackmore, 2002; Dennett, 1991) and conscious will (Wegner, 2002), might be illusions (which nevertheless could have played a significant role in evolution).

- ***Excluded aspects of consciousness.*** To limit the scope of the thesis, I put aside the social sense of *consciousness*, which concerns feelings and beliefs shared by a group, as illustrated by public, national, class, and feminist consciousness. This exclusion is consistent with a focus on psychological rather than sociological explanations, yet it does not apply to social functions of individual consciousness. In part, I also exclude altered states of consciousness. Although dreaming, for example, could have evolved separately (e.g., Panksepp, 1998; Revonsuo, 2000; Snyder, 1966) and may well give hints about normal waking consciousness (Hobson, 2009; Revonsuo, 2006), concentrating on altered states seems an unlikely first approach to the biological evolution and functions of consciousness. In practice, I did not search for means of inquiry in this area (descriptions of their application often fail to mention consciousness), but neither did I ignore them when relevant.

- ***Methodological naturalism.*** Taking a scientific approach to consciousness, including its biological evolution and potential functions, means investigating it in accordance with scientific method (see Haig, 2005, for a recent comprehensive account of scientific method), and thereby seeking natural, as opposed to spiritual or supernatural, explanations of consciousness. Naturalism also extends to means of inquiry (Laudan, 1987; Nickles, 1987b), which are to be evaluated scientifically in terms of their support.

- ***Bases of consciousness.*** Consciousness is generated by biological processes in the brain, which depend on the organism's genetic makeup in combination with its environment, on all relevant time scales. Our conscious experience, for example, typically reflects the relative significance of sight to the human species (Crick & Koch, 1998; Revonsuo, 2006), and also that of interactions with the environment, particularly of a social nature (Donald, 2001; Mead, 1934). Note that the acceptance of this point is not a requirement for the consideration of means of inquiry, or instances of their application, in the thesis; instead, this point helps to explain the position from which I evaluate them.

- ***No consciousness epiphenomenalism.*** Asking about the biological evolution of consciousness would be far less interesting if we adopted epiphenomenalism about consciousness, that is, if we excluded the *possibility* that consciousness has physical effects (see section 1.3). Consciousness could have evolved under epiphenomenalism, but it could not have influenced its possessors' fitness, which I consider an empirical question.

- ***Other conscious minds.*** I shelve the philosophical problem of other minds as it applies to consciousness: How we can know that others have conscious mental states? Such scepticism is effectively ignored by natural scientists (Velmans, 2007), psychologists (Allen & Bekoff, 1997), physicians (Baars, 2005; Marcel, 1988), and, in almost all circumstances, human beings in general (Dennett, 1987; Hyslop, 2010). Section 3.4 comments on the practical problem of determining whether other animals are conscious. In this context it is worth noting that creature consciousness, unlike animal consciousness, does not require the organism to have conscious mental states (Rosenthal, 1986, 2009). The corresponding evolutionary question, namely, why certain organisms are sometimes awake and responsive and at other times dormant or comatose (reviewed by Mignot, 2008; Nicolau, Akaârir, Gamundí, González, & Rial, 2000; Rechtschaffen, 1998; Siegel, 2005), is not dealt with in this thesis.

- ***Conventions.*** For better readability, I often speak of *consciousness* when I mean *consciousness or one or more of its concepts or aspects*. Similarly, the term *the function/s of consciousness* and similar terms should be understood as placeholders for *one or more of the potential functions of consciousness*, allowing for the possibility that consciousness has none. To describe someone or something that is not conscious, whether temporarily or permanently, I prefer the term *unconscious* to *nonconscious*. And, as already noted, I generally use the authors' own terms for consciousness. Moreover, I tend to write about evolution in the past without wanting to imply that it has stopped.

In addition to this list of assumptions and exclusions, my approach can be characterised by pointing to an unusual feature for a theoretical thesis on consciousness. Instead of being concerned with philosophical considerations, it addresses a practical question about scientific research, namely, which means of inquiry should we use if we want to learn more about the evolution of consciousness? Indeed, my role in much of this project is more like that of an engineer who evaluates potential solutions to a design problem than that of a philosopher who argues for a particular position. I believe that the present project is best served not by support for my preferred view but by the basic framework outlined so far, which comprises a research goal and a commitment to science and which aims to include, rather than exclude prematurely, possibly promising means of inquiry in the evaluation.

To conclude this section, our evolutionary question about consciousness is complicated by the fact that consciousness appears to consist of a number of heterogeneous phenomena. It will be crucial to identify, and to find evolutionary explanations for, those (one or more) aspects of consciousness that form evolutionarily relevant units by separating at an appropriate level of description. With a shared understanding of consciousness in place (for the purpose of this thesis), I now turn to essential background material on evolutionary explanations of mind. Both of these first two sections have the same objective: to establish a common denominator in disputed areas of research and thus to prevent misunderstandings in the remainder of the thesis.

1.2 Evolution: A natural approach to mind

When Charles Darwin (1859) laid the foundation of evolutionary biology in his book *On the Origin of Species*, he predicted that "psychology will be based on a new foundation, that of the necessary acquirement of each mental power and capacity by gradation" (p. 488). In the 150 years since then, there have been a very large number of empirical discoveries in evolutionary biology, numerous controversies, and significant theoretical developments such as the evolutionary or modern synthesis. Looking back, Ernst Mayr (2001/2002) commented: "What is most impressive is the robustness of the original Darwinian paradigm" (p. 265). While not everyone agrees with him on the robustness, it is at least the current textbook view (e.g., Freeman & Herron, 2007; Futuyma, 2005; Ridley, 2004). This section begins with a brief review of mechanisms in evolution, then discusses adaptation explanations, and finally explains the evolutionary psychology approach adopted in this thesis.

Evolutionary mechanisms and explanations. Biological evolution can be defined as "*change in the properties of groups of organisms* [populations] *over the course of generations*" (Futuyma, 2005, p. 2). As the main, but not the only, mechanism of such descent with modification, Darwin and Wallace (1858) each proposed natural selection. Natural selection is a statistical process by which a population becomes better adapted, or remains well adapted, to its current environment as follows. Individuals in a population vary in their phenotypes, that is, in the sum totals of their morphological, physiological, behavioural, and other observable properties. Those phenotypes that render their possessors better adapted to the environment are associated with a higher probability of survival and reproductive success (i.e., fitness). Natural selection works by eliminating—through nonconception or early death—some less well adapted individuals in each generation (Broad, 1925; Mayr, 1996). This nonrandom elimination process has an effect on the distribution of properties in future generations only if the advantageous properties are heritable to some degree, that is, if better adapted parents and their offspring tend to resemble each other in the relevant properties. In sum, natural selection results in adaptive evolution when phenotypic variation in a population is tied to differences in fitness and passed on to offspring over many generations.

Other processes besides natural selection are important in evolution, including those that produce the phenotypic variation required for evolution by natural selection. Phenotypic properties under cumulative selection are typically passed on from parent to offspring via the parent's genotype (i.e., its set of genes). Genotypic variation is produced by the process of mutation, which introduces changes in an individual's genetic material, and by recombination of parental genes in the offspring's genotype in sexually reproducing species. Examples of other evolutionary mechanisms include gene flow and genetic drift, that is, movement of genes between populations and random fluctuations in the frequency of alleles (i.e., alternate forms of a gene) within a population, respectively. Environmental factors, of course, also play a major role in evolution—think of geographic differences, interspecies competition, climate change, and natural catastrophes. Taken together, this selection of processes illustrates both the complexity of evolution and the involvement of randomness at many points.

Many properties of organisms are fascinating because of their complexity and remarkable adaptedness for living in the possessor's environment, giving the appearance of purposeful design. Well-known examples include wings for flying, eyes for seeing, echolocation for "seeing" in the dark, camouflage for being less visible, and flowers for attracting pollinators. Darwin (1859) explained such biological adaptations as the result of evolution by natural selection. In fact, "natural selection is the only mechanism known to cause the evolution of adaptations" (Futuyma, 2005, pp. 247–248). Adaptations can accordingly be defined as properties of organisms that have been favoured by nonelimination over many generations because they increased the fitness of their possessors relative to that of nonpossessors. Identifying and explaining adaptations is still a main objective of evolutionary biology (e.g., Alcock, 2005; Freeman & Herron, 2007; Griffiths, 1996; Mayr, 1983; Reeve & Sherman, 1993; Ridley, 2004; Zuk, 2002). The general fascination with adaptations and their special role in evolutionary biology are likely to show in evolutionary hypotheses; and indeed, most evolutionary hypotheses about consciousness do regard it as an adaptation.

While many adaptations are impressive in their adaptedness, "natural selection will not produce absolute perfection" (Darwin, 1859, p. 202). This is because the evolution of adaptations is constrained in many ways. For a start, natural selection makes a generation of organisms better adapted, on average, to the environment in which their parents lived and, hence, is not forward-looking. What is more, natural selection can only work with the variation that exists in a population: A fitter heritable phenotype may not appear for a long time and may then be eliminated by chance. Other genetic constraints stem from, for instance, interactions between genes (i.e., epistasis) and, more generally, the existence of groups of genes. Relatedly, developmental constraints arise from the nature of the developmental system, that is, from the interdependence of processes in an organism's development. Biological structures are, of course, also subject to general physical constraints (e.g., the maximum body size of insects is determined by limits on gas diffusion rates in their respiratory system; Futuyma, 2005). Additionally, trade-offs are often necessary, for example, in the face of competing adaptive demands (e.g., attracting mates vs. avoiding predators) and pleiotropy (i.e., a single gene affecting multiple phenotypic properties). Consciousness should therefore not be assumed to be a perfect adaptive solution, though it could be optimal given the constraints and trade-offs in force.

Because anyone interested in the evolution of consciousness will struggle to sidestep adaptation hypotheses, it is instructive to ask what explanations of adaptations ideally look like. A complete adaptation explanation includes the following components (Brandon, 1990): (a) evidence that natural selection has occurred, for example, from observations in natural populations, experiments, or fossil records; (b) an ecological explanation of better adaptedness, that is, of why natural selection has occurred; (c) evidence for heritability, ideally with reference to genotypes; (d) information about the population structure, that is, about patterns of gene flow and of selective environments; (e) phylogenetic information on trait polarity, that is, on what has evolved from what. However, because the required historical evidence is difficult to come by, "even nearly complete adaptation explanations are going to be rare in evolutionary biology" (Brandon, 1990, p. 177). We have to make do with the most plausible incomplete explanations that agree with general scientific knowledge and hope to move them closer to the ideal by obtaining additional support. Thus, we will not know with a high degree of certainty whether and how consciousness came to be an adaptation, but we might show how evolutionary mechanisms could have produced consciousness.

To recapitulate, what are some of the possibilities to be taken into account when constructing an initial evolutionary explanation of consciousness? First of all, mechanisms other than natural selection, such as genetic drift, gene flow, and environmental factors, could be partially or fully responsible for the evolution of consciousness, which makes it essential to consider such alternate explanations (see Polger & Flanagan, 2002, p. 24, for an admittedly improbable story of its evolution by genetic drift). Further, if consciousness has been tied to another trait, it could be a by-product of that trait's evolution. And although consciousness seems derived from unconsciousness, other scenarios are possible (e.g., Sheets-Johnstone, 1998). If we suspect consciousness to be an adaptation, we should examine its potential heritability and adaptedness, among other things (see Brandon's, 1990, components of adaptation explanation above). Moreover, new adaptations can evolve in different ways: by intensification of a trait's initial function or by a trait's, or a combination of traits', acquisition of a new ecological role.[1] Giving an evolutionary explanation of consciousness is an ambitious goal, with the above possibilities applying to each evolutionarily relevant kind of consciousness in every independent lineage over an extended period of time.

Evolutionary psychology—broadly construed. So far in this section I have simply assumed that evolutionary theory applies to consciousness as it applies to other components of organisms, in line with Darwin's introductory quotation about evolutionary explanations of mind. My main reason for doing so was that biological evolution, and its applicability to all organisms, is nowadays considered a scientific fact (Freeman & Herron, 2007; Futuyma, 2005; Mayr, 2001/2002; National Academy of Sciences & Institute of Medicine, 2008). Many scientists also agree that "our psychological capacities are evolved traits ... subject to natural selection as much ... as our gait,

[1] While the succession of these evolutionary changes is of great interest, additional related terminology such as *exaptation* (Gould & Vrba, 1982) and *spandrel* (Gould & Lewontin, 1979) is confusing (Buss, Haselton, Shackelford, Bleske, & Wakefield, 1998; Dennett, 1995a; Reeve & Sherman, 1993) and not necessary here, as long as we remember that being adaptive is not the same as being an adaptation (Sterelny & Griffiths, 1999).

dentition, or posture" (R. C. Richardson, 2007, p. 9). Although some mental abilities and behaviours have not been around long enough to have been shaped by biological evolution, evolutionary theory does apply to minds: Not only are minds based on evolved brains, minds are crucial for generating behaviours, which make up an important part of phenotypes. I discuss the role of culture in section 1.4 and potential differences between evolutionary explanations of consciousness and other mental traits in section 2.1. Still, only evolution can afford a scientific explanation of the historical emergence and development of consciousness.

Different approaches have been adopted to study human behaviour and the human mind from an evolutionary perspective. The three major approaches currently in use are human behavioural ecology, evolutionary psychology (narrowly construed), and gene–culture coevolution (Gangestad & Simpson, 2007b; Smith, 2000). To study the evolution of consciousness, human behavioural ecologists would require knowledge about consciousness-specific behaviours that vary as a function of the environment, whereas gene–culture coevolutionists would need to know which components of consciousness are transmitted genetically and which culturally. Evolutionary psychologists seem better positioned, not just because their approach is the dominant one (Laland & Brown, 2002), but because they seek to explain psychological mechanisms. Such a cognitive level of explanation fits well with our construct of interest, *consciousness*. While the different approaches should ideally make complementary contributions, evolutionary psychology appears, at least at first, particularly suited for investigating the evolution of consciousness.

The term *evolutionary psychology* has come to designate a specific scientific paradigm within the field of inquiry with the same name (Buller, 2005; hence the qualification *narrowly construed* in the previous paragraph). Although often overlooked, this paradigm–field distinction has been acknowledged by the introduction of new labels for the paradigm such as "inclusive fitness evolutionary psychology" (Caporael, 2001, p. 608), "'Santa Barbara school'" (Laland & Brown, 2002, p. 154), "'Evolutionary Psychology' (capitalized)" (Buller, 2005, p. 12), and "standard evolutionary psychology model" (Moore, 2006, p. 285). Evolutionary psychologists working within the paradigm construct hypotheses about psychological mechanisms in humans as adaptations to human ancestral environments. Key assumptions of the paradigm are thus that the human mind consists of many functionally specialised information-processing devices, and that these mental programs have been designed by natural selection to solve adaptive problems faced by our hunter-gatherer ancestors (see Buss, 1995; Tooby & Cosmides, 1992, 2005). The paradigm has been, on the one hand, advertised as *the* integrative framework for psychology (e.g., Buss, 1995, 2005; Goetz & Shackelford, 2006; Pinker, 2005; Tooby & Cosmides, 1992, 2005) and, on the other hand, heavily criticised (e.g., Buller, 2005; Panksepp & Panksepp, 2000; R. C. Richardson, 2007; Rose & Rose, 2000). The arguments need not be rehearsed here as the above remarks on the paradigm mainly serve to clarify the evolutionary psychology approach from which I distance myself.

Yet this thesis, an evaluation of the promise of means of inquiry for giving an evolutionary explanation of consciousness, is clearly placed within evolutionary psychology—broadly construed. By analogy to the definition of psychology as the study of mind and behaviour, evolutionary psychology, the field of inquiry, can be defined as "the evolutionary study of mind and behavior" (Caporael, 2001, p. 608), that is, "the study of how mind and behavior have evolved" (Heyes, 2000, p. 3). Means of inquiry in this thesis and the resulting evolutionary explanations are thus not bound by the controversial assumptions of the evolutionary psychology paradigm. The paradigm's focus on human ancestry, for instance, would be limiting for the following reason:

> How did human consciousness evolve? This is a question that psychologists love to ask. The answer is actually quite simple: from animal consciousness! ... it is quite certain that human consciousness did not arise full-fledged with the human species, but is only the most highly evolved end point of a long evolutionary history. (Mayr, 2001/2002, p. 282)

We therefore need a more complete "evolutionary psychology 'in the round'" (Heyes, 2000, p. 3) for investigating the evolution of consciousness in humans and other organisms.

The present section heading called evolution, or rather evolutionary studies, *a natural approach to mind.* Evolution is, of course, an important part of the natural world, and so is the mind, at least as studied scientifically. Asking where such a natural phenomenon came from and how it was shaped by natural processes over time seems a *natural* question—in several senses of the word. For example, the evolutionary question seems more natural than, say, questions relating mind to computer metaphors, artificial intelligence, or current Western culture. Answers to other natural questions, such as how the brain generates the mind, will strengthen the evolutionary account of mind by adding knowledge at a different level of explanation. Here I have provided a brief evolutionary background in order to give a taste of what might be involved in an evolutionary explanation of consciousness, the research goal in the evaluation of means of inquiry. It is neither necessary to restrict these means to a single evolutionary approach, nor is it wise to do so. In the remainder of the chapter, I discuss objections to evolutionary explanations of consciousness that consider such explanations insignificant or even impossible, not just difficult.

1.3 Epiphenomenalist worries addressed

What if it could be shown that consciousness has no influence whatsoever on any part of the physical world? If consciousness had no such effects, then we would know, for instance, that consciousness does not affect the fitness of its possessors. We could still ask how it evolved, but not find out. Consequently, if epiphenomenalism about consciousness were true, there would be no point in evaluating means of inquiry into the biological evolution of consciousness. Somewhat reminiscent of the significance attributed to unconscious mental processes in the study of consciousness (e.g., Baars, 2003b; Frith & Rees, 2007; Güzeldere, 1995a), Humphrey (1987a) suggested that "the realization that consciousness might be useless was ... something of a breakthrough" (p. 378). However, he continued, "it is a naughty idea which has, I think, had a good run, and now should be dismissed" (p. 378). Let us see whether this thesis-endangering possibility—that consciousness makes no difference at all—can be dismissed.

The easiest way to make consciousness epiphenomenal is by defining it so. The object of this manoeuvre is typically phenomenal consciousness in one form or another. For example, Chalmers (1997) argued that the strong empirical correlation between phenomenal and access consciousness means that any function can be ascribed to the latter. A similar move has been performed in an evolutionary context:

> So does consciousness in this third sense of sentience ... have an adaptive function? ... I think the answer has to be no, simply because it's defined in such a way that it has no causal consequences ... Anything that does have causal consequences we peel off or distill out and assimilate to the ... information access problem of consciousness. (Pinker, 2004)

Defining consciousness as incapable of influencing anything (sometimes with the exception of other epiphenomenal mental events) is effectively defining it as not scientifically investigable and, hence, violates methodological naturalism. Allowing consciousness to have physical effects but defining it as functionless still pre-empts the empirical investigation of its function/s. Evolutionary studies of consciousness best avoid epiphenomenal *definitions* of consciousness altogether.

As an aside, consciousness epiphenomenalism is fostered by a highly influential intuition, termed the *segregationist intuition* (Güzeldere, 1997; *essentialist intuition* in Güzeldere, 1995a, 1995b). According to this intuition, the following two pretheoretic characterisations of consciousness are mutually exclusive: "'Consciousness is as consciousness *does*' versus 'Consciousness is as consciousness *seems*'" (Güzeldere, 1997, p. 11). This apparent dichotomy between the causal and the phenomenal characterisation of consciousness is closely related to the third- versus first-person perspective, and underlies such distinctions as access versus phenomenal consciousness and easy versus hard problems of consciousness, as well as several other debates (see Güzeldere, 1997). In keeping with the segregationist intuition, epiphenomenalists consider consciousness as fundamentally phenomenal and therefore essentially noncausal. This intuition needs to be replaced by a more integrative view, as Güzeldere (1997) argued, in order to resolve many of the debates on consciousness.

Varieties of epiphenomenalism. Among the varieties of consciousness epiphenomenalism to be differentiated, *strict metaphysical epiphenomenalism* (Polger & Flanagan, 2002) is the standard philosophical sense of the term already introduced: While mental events are caused by physical events, "no mental event plays any part in the causation of any bodily event" (Broad, 1925, p. 118). This variety of epiphenomenalism originated when mental properties were generally believed to be nonphysical, a kind of mind–body dualism. Its rationale is that, given dualism, mental events *cannot* have any physical effects because every physical event has a sufficient physical cause and is not overdetermined (Polger & Flanagan, 2002; Robb & Heil, 2009; Robinson, 2009; cf. "the Master Argument for Consciousness Epiphenomenalism," Graham, 1998, p. 229). An overview of some of the theoretical and empirical support for and against epiphenomenalism is given in the next subsection. Note that in speaking of *mental events* I have presented the most general version of strict metaphysical epiphenomenalism; such epiphenomenalism about consciousness, or one or more of its concepts or aspects, is of more concern here.

A second variety of epiphenomenalism about consciousness, called *causal role epi-*

phenomenalism (Polger & Flanagan, 2002), allows consciousness to have physical effects, but no causal role function in conscious organisms. Something has a causal role function if it is capable of producing effects that help explain a capacity of its containing system (see Cummins, 1975, p. 762, for a formal statement). As regards consciousness, a particular capacity of conscious organisms or of one of their components (e.g., fitness; Buller, 1998) might be explained, in part, by the capacity of consciousness to produce certain effects (or, if seen as an activity or behaviour, by the organism's capacity to engage in consciousness). Causal role epiphenomenalists deny that consciousness has any such important effects, thereby using the term *epiphenomenon* more or less in its original medical sense of *secondary symptom.* To illustrate, the most famous analogy of epiphenomenal consciousness (actually given for strict metaphysical epiphenomenalism; N. Campbell, 2001; Polger & Flanagan, 2002) compares consciousness to "the steam-whistle which accompanies the work of a locomotive engine [and which] is without influence upon its machinery" (Huxley, 1874/1882, p. 236). Causal role epiphenomenalists about consciousness thus claim that none of the effects that consciousness has contribute causally to any of the capacities of conscious organisms.

The least restrictive variety of epiphenomenalism is called *etiological epiphenomenalism* (Polger & Flanagan, 2002; termed *biological epiphenomenalism* by Revonsuo, 2006). Etiological epiphenomenalists about consciousness maintain only that consciousness has no etiological function, that is, that consciousness is not a biological adaptation (historically defined, as in the previous section; Amundson & Lauder, 1994). The etiological function of a trait can be thought of as a causal role function for which the trait has been naturally selected over evolutionary time (cf. Buller, 1998; Griffiths, 1993). Despite differences in generality between etiological and causal role functions, their distinction brings to mind Mayr's (1961) distinction between evolutionary and functional biology and Tinbergen's (1963) distinction of biological questions about evolution, survival value, causation, and ontogeny. While proximate explanations will be needed for a complete explanation of why we are conscious, the overall research goal here is ultimate explanations of consciousness. Etiological epiphenomenalism about consciousness is directly relevant to this goal and opposite to the research strategy evaluated in Chapter 3, which aims to show that consciousness is an adaptation.

Pros and cons, theoretically oriented. To start looking at arguments for and against epiphenomenalism, let an intriguing creature enter the scene: the philosophical zombie. Philosophical zombies are imaginary creatures that are behaviourally identical to human beings—many of them are also functionally or even physically identical to us (Güzeldere, 1995c; Polger, 2000b)—but they lack consciousness, nowadays commonly phenomenal consciousness. The contentious issue about philosophical zombies is whether they are not only logically but metaphysically or even naturally possible (Flanagan & Polger, 1995; see Polger, 2000a, for a systematic overview of identity vs. possibility combinations). Some researchers have answered this question based on their preferred philosophical theory of consciousness in order to explore the theory's consequences (Polger, 2000a). More often, the conceivability or inconceivability of zombies is used as an independent argument for or against such a theory (see Kirk, 2009, for a collection of arguments about their conceivability and inferred possibility). To give an early example, Stout (1931) argued against epiphenomenalism that "there can be no doubt that this [physically identical zombie world] is *primâ facie* incredible to Common Sense [the convergence of lay and expert views]" (p. 139). If such a physical zombie world were naturally possible, our consciousness would not make any (physical) difference. But this possibility conflicts with my assumption, shared by epiphenomenalists, that consciousness is generated by neural processes in accordance with causal laws (see section 1.1), whether in our or in physically identical zombie brains. We can thus at least put aside the threat of strict metaphysical epiphenomenalism from physically identical zombies.

As the following theoretical arguments for epiphenomenalism indicate, it is typically only defended when all alternatives appear even less attractive (Pauen, Staudacher, & Walter, 2006). The first argument has already been sketched as the rationale for strict metaphysical epiphenomenalism: The causal irrelevance of the mental in the physical world results from considering the mental—whether its substance or some of its properties—as radically different from the physical, while accepting that all physical events can be explained by physical causes alone. Several other difficulties for explaining how mental events *could* play a causal role in behaviour also lend support to epiphenomenalism, including externalism (i.e., the theory that mental content depends on conditions external to the individual) and a denial of appropriate mental–physical laws (see Robinson, 2009, and Walter, 2007, for summaries). Furthermore, epiphenomenalists can argue that it is more plausible that the complex neural events that cause sensations also cause the complex neural events required for behaviour than that the latter events are caused by simple sensations (Robinson, 2009). Although it seems to us that our mental life matters for what we do, arguments for epiphenomenalism suggest that it is far from clear that this is the case.

To say that "epiphenomenalism has had few friends" (Walter, 2007, section 5, para. 1) is an understatement, but have opponents been able to disarm it? (See Robinson, 2009, and Walter, 2007, for summaries of the following arguments and replies.) What may be the most persuasive argument, namely that epiphenomenalism is counterintuitive if not absurd, does not detract from its potential truth. Nor does naively insisting on introspective evidence work, because introspection does not allow us to distinguish whether the mental event itself or its physical cause led to the internally or externally observed effect. Further, epiphenomenalists can infer the existence of other minds from similar behaviour, not via assumed similar mental causes of behaviour, but via assumed similar physical causes of behaviour and mental events. They can similarly argue that natural selection could not have favoured the mind but its neural causes and that the mind would then have evolved as a by-product of these neural causes of fitness-enhancing behaviour. The most powerful argument against epiphenomenalism questions how epiphenomenalists can know of mental events if these events have no physical effects and, hence, cannot cause such knowledge. Epiphenomenalists might reply, for example, that knowledge about mental events is caused by their physical causes or that we are directly acquainted with our experiences. As this selection of arguments against epiphenomenalism shows, the theory is fairly resistant to theoretical attacks.

Pros and cons, empirically oriented. Empirical evidence bearing on epiphenomenalism about consciousness normally concerns whether consciousness has any causal role function, not whether it is a biological adaptation or has any physical effects at all. Many researchers have deemed a study of voluntary action by Libet, Gleason, Wright, and Pearl (1983) "An Experiment in Epiphenomenalism" (Flanagan, 1992, p. 136). In this classic study, participants reported becoming aware of their intention to spontaneously perform a given motor act on average 343 ms after motor areas in their brains had already started preparing the movement. The perceived epiphenomenalist threat derives from, in the words of the study's title, "The Unconscious Initiation of a Freely Voluntary Act."[2] However, the finding that consciousness of

[2]The study has been widely criticised (e.g., commentaries following Libet, 1985), but also replicated (Haggard & Eimer, 1999; Keller & Heckhausen, 1990; Sirigu et al., 2004; Trevena & Miller, 2002); and

the intention to initiate an action followed neural preparations should not surprise us if we assume that consciousness emerges from brain processes. More importantly, it does not mean that consciousness makes no difference to behaviour overall.

A more comprehensive defence of consciousness epiphenomenalism was presented by Velmans (1991). Based on a review of the experimental literature on various cognitive domains, he concluded that consciousness is epiphenomenal (from a third-person perspective) because it performs none of the many functions claimed for it. One reason for this is that cognitive processes such as perception and learning can occur unconsciously; another that we frequently become aware only of the results of mental processing. To explain why consciousness accompanies some mental processes some of the time and even appears necessary for them, Velmans suggested that "consciousness nearly always *results* from focal-attentive processing" (p. 651), which is what influences subsequent processing. Note, in passing, that this reflects the segregationist intuition: Causal and noncausal aspects are being kept separate by having focal-attentive processing do all the work. Given the numerous points of criticism raised (e.g., in the commentaries following Velmans, 1991), to which I cannot do justice here, Velmans's conclusion should be regarded as a controversial hypothesis for further research.

Other empirical evidence that has been enlisted as support for or against epiphenomenalism comes from the structure and neural bases of consciousness and from neurological deficits. According to Flanagan (1992), "the biggest problem the epiphenomenalist faces is explaining how ... *any* feature as common, well-structured, and multimodal as phenomenal consciousness could ... [not have] interesting and important causal effects in other parts of the neural network" (p. 150). Because the neural bases of consciousness are distributed widely in a highly connected brain, they are likely to affect subsequent processing and to endow consciousness with causal role function/s (Baars, 2003a; Flanagan, 1992). Moreover, neurological deficits in which some of these neural bases are disrupted tend to be associated with more or less severe changes in psychological functioning and behaviour. Patients with blindsight, a well-known example, report seeing nothing in areas of blindness caused by damage to their primary visual cortices, yet they are able to respond to certain visual stimuli presented to their blind fields. This ability has been taken to support the

the conscious experience of volition has become an active research area in neuroscience (see Haggard, 2005, 2008, and Hallett, 2007, for reviews).

insignificance (P. S. Churchland, 1983) and causal irrelevance (Velmans, 1991) of consciousness. However, blindsight patients do not normally initiate voluntary actions based on visual information from their blind fields, indicating a potential function of consciousness (e.g., Flanagan, 1992; Marcel, 1986, 1988; Van Gulick, 1989; but see Block, 1995; de Gelder et al., 2008; Stoerig, 2010). Other means of inquiry for identifying the (evolutionary or nonevolutionary) function/s of consciousness, evaluated in Chapters 4 and 5, may help to make further empirical inroads into the area of epiphenomenalist contention.

In conclusion, epiphenomenalism about consciousness cannot be easily dismissed (Flanagan, 1992; Pauen et al., 2006; A. K. Seth, 2009; Van Gulick, 2009). Intuitions and metaphysical commitments often determine the position taken on any of the varieties of epiphenomenalism. While metaphysical epiphenomenalism about consciousness is (probably) not empirically testable (Pockett, 2004), scientists could settle for an explanation of the evolution of the neural causes of consciousness and their effects. Note that I assumed *no consciousness epiphenomenalism* in section 1.1 merely to avoid a premature exclusion of the possibility that consciousness has physical effects. Causal role and etiological epiphenomenalism about consciousness should be treated as empirical questions (Flanagan, 1992); examples of relevant studies have been given above, and "there is plenty of other evidence relevant to EP [epiphenomenalism]" (A. K. Seth, 2009, p. 282). In so far as these two varieties of epiphenomenalism matter to the evaluation of means of inquiry, they are taken up again in Chapter 3 on the adaptation status of consciousness and Chapters 4 and 5 on its evolutionary function/s and history. In addition, we reencounter zombies when we consider three other potential obstacles to evolutionary explanations of consciousness in the next subsection. Ultimately, and in agreement with Flanagan (1992), epiphenomenalism is a healthy reminder that we cannot take for granted that consciousness is functional and important.

1.4 Other doubts about the approach dismissed

There are, of course, further reasons for disagreeing with the suggestion that "the phenomena of consciousness may best be approached through examining its evolutionary significance" (Bechtel & Richardson, 1983, p. 389). First of all, there are other important approaches to consciousness, and there is no need to claim primacy of the evolutionary approach when the different approaches can, in fact, usefully complement each other. The evolutionary approach is certainly not the most straightforward approach to consciousness and has frequently misled theorising about its nature (Polger, 2007). Before examining difficulties of its application in the next chapter, it is worth dealing with the following objections: First, consciousness is not necessary for any of our mental abilities or behaviours because they can in principle be carried out unconsciously. Second, biological evolution is irrelevant to explanations of consciousness because consciousness is too recent a product of culture. Third, it is impossible for humans to understand consciousness due to our cognitive limitations. Addressing these objections in turn, I argue that none of them foredoom the current project to failure.

Conscious inessentialism. Conscious inessentialism is the view that any mental activity can in principle be performed unconsciously, even if it is accompanied by consciousness in us (Flanagan, 1991, 1992). This view of consciousness as inessential to intelligent mentality became plausible with the emergence and development of cognitivism and artificial intelligence. It is expressed in statements such as Fodor's (2004) that, "as far as anybody knows, anything that our conscious minds do they could do just as well if they weren't conscious" (p. 31). And here we already reencounter philosophical zombies: If everything we do could in principle be done without consciousness, behavioural zombies are logically and metaphysically possible (Flanagan, 1991; Flanagan & Polger, 1995). Thus, in the most commonly used words, conscious inessentialists hold that consciousness is simply *not necessary* for any of our mental abilities or behaviours.

Conscious inessentialism has worried many consciousness researchers because it seems to imply epiphenomenalism. For example, Dennett (1987) described the worry created by cognitive psychologists' theorising about unconscious mental processes by asking, "What is consciousness *for*, if perfectly unconscious, indeed subject-less, information processing is in principle capable of achieving all the ends for which conscious minds were supposed to exist?" (p. 162). Similar reasoning has been employed in an evolutionary context (e.g., Blackmore, 2001; Hameroff, 1999; O'Hear, 1997). In fact, both zombie arguments and some of the empirical arguments for epiphenomenalism in the previous section (e.g., Velmans's, 1991, inferences from implicit learning and blindsight) reason from the assertion that consciousness is not required for particular abilities to the conclusion that consciousness is not causally involved in them.

However, epiphenomenalism does not follow from inessentialism: Even if there are unconscious ways of producing our behaviours, consciousness—when present—could still affect their generation in us (Dretske, 1997; Polger, 2007; Searle, 1998; A. K. Seth, 2009). (The empirical arguments, which are not deductive, tend to suffer more from differences between the compared conscious and unconscious behaviours.) In general, then, conscious inessentialism gives much less reason than commonly assumed to question potential causal role function/s of consciousness in us.

When claiming that consciousness is necessary (or not necessary) for a particular ability, researchers may have different kinds of necessity in mind. On the one hand, there is "the deep metaphysical sense involving possible beings in possible worlds" (Flanagan, 1991, p. 344). An ability for which consciousness is held to be strictly necessary cannot be shown by anyone or anything in any way that is not conscious. This kind of necessity is reflected in Humphrey's (1987a) challenge: "If that ability [of using self-knowledge to interpret others] could exist without consciousness, let someone prove it to me" (p. 381). On the other hand, consciousness might only be necessary for us—as we have come to be in the actual world—to exercise a particular ability well under certain conditions. It is such contingent necessity in organisms that may call for an evolutionary explanation, precisely because organisms' contingent properties did not have to arise (Polger, 2007; Polger & Flanagan, 2002; Williams, 1966). When considering suggested function/s of consciousness, we need to be clear whether consciousness is claimed to be strictly, contingently, or not necessary for a particular ability.

The claimed necessity requires a proximate explanation, and it affects what evolutionary explanations of consciousness need to account for. If consciousness is strictly necessary for an evolutionarily significant ability, contrary to conscious inessentialism, then the evolutionary history of consciousness is (most likely) the same as the ability's (Polger, 2002), at least since the time consciousness became necessary for that ability. So to explain the evolution of consciousness, we would want an evolutionary explanation of the ability for which it is strictly necessary (cf. Blackmore, 2001). Yet strict necessity is an unlikely claim for biological traits (Polger, 2007; A. K. Seth, 2009), unless one adopts the functionalist view that whatever does what consciousness does *is* consciousness. If consciousness is instead only contingently necessary for an ability in us, consistent with conscious inessentialism, then we need to explain not only how the ability evolved but also why consciousness, rather than some other process, came to contribute to the ability in our ancestors (Polger, 2002). How current proposals for the evolution of consciousness fare in providing appropriate evolutionary explanations is a subject of the next chapter.

In brief, conscious inessentialism does not imply epiphenomenalism, but it does underscore the fact that "consciousness did not have to evolve" (Flanagan, 1991, p. 344). This is why, for example, an explanation of what consciousness is good for should be supplemented by an explanation of why consciousness became associated with this benefit. Flanagan and Polger (1995) described the worth of inessentialism for consciousness research as follows:

> Recognition that consciousness did not have to be highlights the pressing questions of why it exists, why it evolved, whether there were competing zombie hominids who lost out in the struggle to survive, and if there were not why not? (p. 321)

I suspect that, although consciousness may not be strictly necessary for an ability, it was the only or best way to achieve that ability when it first arose in our ancestors. In any case, both epiphenomenalism and inessentialism serve to draw our attention to problematic assumptions about the function/s and evolution of consciousness.

Culture versus biology. Before looking at two theories of consciousness according to which it is a cultural product that has not been shaped by biological evolution, the relationships between culture, biology, and evolution deserve some comments. Biology and culture have at times been taken to furnish separate or even conflicting explanations of human mind and behaviour, as in the nature–nurture debate about determinants of human development. However, biological and cultural explanations are often complementary, not mutually exclusive. For example, when researchers describe consciousness either as a product of biology or as one of culture, they are likely to be pursuing different research questions about different aspects of consciousness. Besides, culture itself is a biological phenomenon (Alcock, 2005; Buller, 2005; Sterelny & Griffiths, 1999; Tooby & Cosmides, 1992), in the following sense: Human cultures are created by biological creatures whose cultural behaviours "differ only in degree of complexity, not in kind" (Buller, 2005, p. 422), from behaviours of other animals (Byrne et al., 2004; Laland & Hoppitt, 2003; Whiten, Horner, & Marshall-Pescini, 2003). Biological (again in a somewhat narrower sense) and cultural explanations are, on the whole, more compatible than often portrayed.

Turning now to evolution, cultural transmission and biological evolution are distinct, but they do influence each other. How far the analogy between biological and cultural evolution goes—a contentious issue (Henrich, Boyd, & Richerson, 2008; Henrich & McElreath, 2003; Mesoudi, 2007; Mesoudi, Whiten, & Laland, 2006)—is not central to the interdependence of biological and cultural processes, which is of interest here. To begin with, our brains have evolved to be highly structured as well as highly plastic (Gopnik, 2010). The specifics of our so-evolved minds enable culture and its transmission on the one hand, and guide and constrain them on the other (Laland, Odling-Smee, & Feldman, 2000; Tooby & Cosmides, 1992). Cultural activities in turn affect biological evolution by changing the diversity of phenotypes on which natural selection acts and by modifying selective environments (Laland, Kendal, & Brown, 2007; Laland et al., 2000). The idea that "humans have coevolved with their culture" (Sterelny & Griffiths, 1999, p. 19) is also behind the gene–culture coevolutionary approach to human behaviour. Biological and cultural evolution can interact, with the details of the coevolution depending on the phenomenon under investigation.

Consciousness was invented by humans as recently as about 1000 B.C., according to Jaynes's (1976/2000, 1986) theory of the origin of consciousness. Jaynes (1976/2000) inferred from archæological artefacts and literary texts such as the *Iliad*, which generally lacks terms for consciousness and mental acts, that conscious minds were preceded by a very different mentality based on hallucinatory voices giving directions. Jaynes called these earlier minds *bicameral* to indicate their division into an executive and a follower part. After bicamerality had broken down for various reasons, consciousness emerged as a new method for deciding what to do. In Jaynes's theory, consciousness is "a metaphor-generated model of the world" (p. 66) that is "learned on the basis of language and taught to others" (p. 220). As its origin he suggested the positing of internal causes of behaviour in strangers to explain the bewildering difference of their behaviours. Jaynes emphasised that "consciousness is chiefly a cultural introduction" (p. 220), but allowed for "a modicum of natural selection" (p. 221) during the upheavals in the later centuries of the second millennium B.C., when less conscious individuals were more habitual or impulsive and incapable of long-term deceit. Nonetheless, consciousness is pictured mainly as a cultural product.

Our mentality is undoubtedly influenced strongly by social learning and language, yet other aspects of Jaynes's (1976/2000) theory are less secure. Here I concentrate on the two most pertinent points. First, being conscious in Jaynes's sense means having the concept of consciousness, as Dennett (1986) explained and Jaynes confirmed. We could thus accept Jaynes's evolutionary hypothesis for this conceptual kind of consciousness while disagreeing that it exhausts all of consciousness. Relatedly, consciousness has been described as a model of the world by others too (e.g., Merker, 2005; Metzinger, 2003; Revonsuo, 2006), but is it really generated through language alone? Second, Jaynes's theory seems to rest largely upon reasoning from inessentialism to epiphenomenalism, though neither of them in full-blown form. Jaynes concluded from a review like Velmans's (1991) that consciousness is not necessary for many of our activities and urged the possibility of "human beings who did most of the things we do—speak, understand, perceive, solve problems—but who were without consciousness" (i.e., our near-zombie ancestors with bicameral minds; Jaynes, 1986, p. 131). Because this thesis is about the biological evolution of consciousness, only Jaynes's (1976/2000) means for constructing the natural-selection hypothesis are relevant, not those for studying the cultural evolution of consciousness. Studying cultural changes in consciousness over the past few millennia, albeit not the focus here, is another interesting project.

Similar hypotheses about the recency and culture-dependence of consciousness have been advanced by Dennett (1991): "Human consciousness (1) is too recent an innovation to be hard-wired into the innate machinery, [and] (2) is largely a product of cultural evolution that gets imparted to brains in early training" (p. 219). The first point is shaky in light of evidence that natural selection can cause significant genetic change in a short time (e.g., S. P. Carroll, Hendry, Reznick, & Fox, 2007; Kingsolver & Pfennig, 2007). The second point may be granted for some aspects of consciousness, including the aforementioned conceptual consciousness, but not for other features and kinds of consciousness such as phenomenal and access consciousness (Block, 1993, 1995; Flanagan, 1992). Dennett's proposal, including a third point, has been labelled "*very* contentious" (Flanagan, 1992, p. 84). The disagreement appears to stem, in part, from differences in emphasis of primary versus language-dependent aspects of consciousness. In the spirit of the latter—but without completely neglecting the former—Dennett stated that "the sort of consciousness such [languageless] animals enjoy is dramatically truncated, compared to ours" (p. 447), which has been shaped by both cultural and biological evolution. Human consciousness *is* a cultural product to some degree; yet this is not in conflict with the research goal against which means of inquiry are evaluated here, namely, discovering the biological origin and history of consciousness.

Cognitive closure. The final threat to the present project addressed in this chapter is the new-mysterian view that "consciousness, despite being a natural phenomenon, will *never* be understood" (Flanagan, 1992, p. 9). Instead of blaming the nature of consciousness for the topic's perceived difficulty, new mysterians claim that our cognitive limitations are at fault. In McGinn's (1991) words, it is "our own incurable cognitive poverty" (p. 43) that precludes us in principle from explaining how consciousness arises from the brain, forever leaving us with "a vertiginous sense of ultimate mystery" (p. 7). In the concluding paragraph of his *How the Mind Works*, Pinker (1997) commented that "if these conjectures are correct, our psyche would present us with the ultimate tease. The most undeniable thing there is, our own awareness, would be forever beyond our conceptual grasp" (p. 565). In short, new mysterians say that human minds are *cognitively closed* with respect to the link between consciousness and the brain—we cannot hope to ever achieve "closure" on consciousness.

After a chapter-long critique of McGinn's (1989, 1991) main arguments for new mysterianism, Flanagan (1992) concluded that, although the view is coherent, the arguments for it are unconvincing. To give an example not discussed by Flanagan, new mysterianism is motivated by the persistent failure to solve the mind–body problem (McGinn, 1989; Pinker, 1997); yet ignorance itself is not particularly informative (P. S. Churchland, 1996), and scientific progress has made many formerly mysterious phenomena much better understood (P. S. Churchland, 1988; Polger & Flanagan, 1999; Searle, 1993, 1998). Possibly the strongest objection to new mysterianism is that we should not set higher standards on (scientific) explanations of consciousness than on other explanations by requiring the former to "provide a *fully satisfying* [emphasis added] account of consciousness" (McGinn, 1991, p. 28), "a solution that feels satisfactory above and beyond ... [being] true" (Pinker, 2004; see Flanagan, 1992; Polger & Flanagan, 1999; Vaneechoutte, 2000; Wright, 2007). The apparent irreconcilability of first- and third-person perspectives drives McGinn's (1989) argument for new mysterianism (Flanagan, 1992), which is yet another sign of the segregationist intuition (Güzeldere, 1995b). Until there is much stronger support for cognitive closure, I think it is best to treat it as an empirical hypothesis, as in fact Pinker (1997, 2004) does. I therefore endorse Flanagan's recommendation to get on with research rather than to rely on new mysterianism being accurate.

Let us consider what cognitive closure, if found to be true, would mean for evolutionary explanations of consciousness. To start with, the computational and neural aspects of consciousness and their evolution are "perfectly tractable" (Pinker, 1997, p. 563) under new mysterianism. As McGinn (1989) acknowledged, "the brain has physical properties we can grasp, and variations in these correlate with changes in consciousness" (p. 364). However, he maintained that we will never be able to answer the question "How did evolution convert the water of biological tissue into the wine of consciousness?" (McGinn, 1999, p. 13). That is, any account of the evolution of consciousness will be an account of the evolution of its neural bases, without a full explanation of how these bases create consciousness. Carruthers's (2000) conclusion that such accounts will not be "accounts of the evolution of [phenomenal] p-consciousness *as such*" (p. 256) is too harsh because evolutionary explanations address ultimate, not proximate, questions. New mysterianism denies (us) the possibility of a complete explanation of consciousness, but not that of its evolutionary explanation.

None of the three objections in this section deal a fatal blow to evolutionary explanations of consciousness; rather, they contribute to the present project by pointing to areas in need of research attention. First, conscious inessentialism, which does not entail epiphenomenalism, made us realise that care is required when consciousness is said to be necessary for an ability: Strict necessity makes evolutionary explanations redundant, whereas contingent necessity calls for the consideration of alternatives. Second, cultural influences on the evolution of human consciousness, though interesting, are outside the scope of the present project, as opposed to the biological origins of these cultural aspects of consciousness. Third, cognitive closure as advanced by new mysterians is an empirical possibility, which does not directly affect evolutionary explanations of consciousness. We have also seen that we should not ask too much of explanations, such as proximate explanations of evolutionary accounts or satisfaction. In this section we have cleared some of the ground for evolutionary explanations of consciousness by removing alleged obstacles and clarifying the present project.

Conclusion

This chapter introduced the two core components of the research goal against which means of inquiry are evaluated in subsequent chapters: consciousness and biological evolution. Starting from a basic understanding of consciousness as an organism's awareness of the environment and of itself in it, we found that an essential task in the evolutionary study of consciousness is determining its evolutionarily relevant units. A brief sketch of evolutionary theory showed that researchers who are constructing evolutionary explanations of consciousness should consider details of the trait's emergence and development over time, and other mechanisms besides natural selection. Means for investigating one such potential factor in the evolution of human consciousness in particular, namely culture, are not included in the evaluation. This first chapter presented the minimal knowledge of consciousness and evolution needed to appreciate the contribution that means of inquiry make toward the research goal of giving a scientifically respectable evolutionary explanation of consciousness.

The chapter began with a look at "consciousness ... the biggest mystery" (Chalmers, 1996, p. xi), a theme that reappeared in later sections. New mysterians, for example, take the long-standing perceived mystery of the connection between consciousness and the brain as indicating that human minds will never understand it. Such cognitive closure would concern the proximate question of how the brain generates consciousness, not the ultimate question of how consciousness evolved. Another reason "why consciousness is so baffling ... [is that] it seems to be among the chronically unemployed" (Fodor, 2004, p. 31). This brings us back to epiphenomenalism and inessentialism about consciousness, which make plain that we cannot simply assume consciousness to be functional and necessary in principle for any ability. The seeming mystery and confusion about consciousness, both the phenomenon and its concept, result to a considerable extent from the segregationist intuition, which takes causal and phenomenal characterisations of consciousness to be mutually exclusive (Güzeldere, 1997). This diagnosis does not instantly remove the mystery, but it offers an explanation of it and suggests how to approach the problem.

Now that the groundwork has been laid—the two main content areas have been introduced and four worries about the evolutionary approach to consciousness have been eased—we can move on: "Once one adopts the view that consciousness is a natural process that occurs in some kind of creatures, then there is no philosophical puzzle about how consciousness evolved, just the hard work of evolutionary biology" (Polger, 2007, p. 82). And this hard work requires means of inquiry, the most promising ones we have. Note that in order to be able to include many potentially useful means in the evaluation, I have imposed as few limitations on them as I deem reasonable within the scope of the topic. For instance, means of inquiry do not need to be aligned with narrowly-construed evolutionary psychology. Furthermore, the thesis takes a scientific realist perspective and only expects means to contribute to a natural explanation of consciousness in accordance with scientific method. The next chapter presents my evaluation tool for means of inquiry into the biological evolution of consciousness, after explaining why the evaluation is timely and important.

2 A tool for evaluating the promise of means of inquiry

Having a better idea of what we want—a scientifically respectable evolutionary explanation of consciousness—does not count for much if we do not know how to get there. Finding one or more good ways to the research goal is, in turn, likely to be helped by a better understanding of where we are in relation to the research goal. This chapter therefore begins with an examination of the current state of the research literature on the biological evolution of consciousness, where indications of difficulties in the scientific study of consciousness and its evolution soon become apparent. Yet is consciousness really particularly difficult to study? And is explaining the evolutionary origin and function/s of consciousness "the hardest problem in consciousness studies" (Flanagan & Polger, 1995, p. 313)? Once the problems are clarified as methodological and mostly not unique, I suggest enabling their solution by evaluating the promise of means of inquiry in the present context. To be able to implement this suggestion in the following chapters, I present here an evaluation tool constructed for this purpose.

The first section of the chapter contains the aforementioned characterisation of the current state of evolutionary explanations of consciousness, followed by an analysis of the perceived difficulties in studying consciousness. I then explain what benefits an evaluation of means of inquiry could have as well as why I needed to construct a suitable evaluation tool. The first half of the chapter thus ends the introductory materials by providing a rationale for the project. The second half of the chapter forms the method section of the thesis. After describing the considerations and steps involved in the construction of the evaluation tool, I introduce the tool for evaluating the promise of means of inquiry that have been used in or suggested for the study of the evolution and function/s of consciousness. The tool generates a promise score for a means by combining weighted criteria scores on three dimensions, which indicate the means's relevance, efficacy, and practicality.

2.1 Current explanations and perceived difficulties

Almost two decades ago, the following characterisation of the literature on the biological evolution of consciousness was offered by Dennett (1991): "There have been many theories—well, speculations—about the evolution of human consciousness, beginning with Darwin's own surmises in *The Descent of Man* (1871)" (p. 172). Biological adaptations have in fact been studied since long before Darwin (Mayr, 1983), but factors such as the prevalence of mind–body dualism (Crook, 1980) and behaviourism (Mandler, 1975) impeded the evolutionary study of consciousness. Given the growing interest in consciousness as a research topic in the last two decades, it is worth inquiring how evolutionary explanations of consciousness have developed since Dennett's description: Is his characterisation of the literature on the evolution of consciousness as consisting of many speculations accurate today? And if it is, are the reasons for this state of the literature inherent in the study of consciousness and its evolution, and possibly even insurmountable?

State of evolutionary and functional explanations. The literature on the evolution and function/s of consciousness and, by inference, research on this topic are in transition, as illustrated by the following assessments and examples. In spite of "the contemporary fuss and feathers about the nature and function of consciousness" (Mandler, 2002, p. vii), "on the whole, there is much about the ... question [of the function and evolution of consciousness] that remains to be written than what is already there" (Güzeldere, 1997, p. 36). The relatively recent renewal of interest in the topic and the unsatisfactory—but improving—state of the corresponding literature are also reflected, for instance, in evolutionary psychology publications: Consciousness has clearly not been part of mainstream evolutionary psychology (Bering & Bjorklund, 2007), yet several evolutionary psychology textbooks do cover it in some detail (e.g., Bridgeman, 2003; Hampton, 2009), and increasingly so (e.g., Gaulin & McBurney, 2001, 2004; Workman & Reader, 2004, 2008). In addition to a higher publication rate of journal articles on the topic (e.g., Cabanac et al., 2009; Ginsburg & Jablonka, 2007a, 2007b; Merker, 2005; Morsella, 2005), edited collections covering the evolution (Fetzer, 2002; Liljenström & Århem, 2008; Terrace & Metcalfe, 2005) and function/s (Ganten, Gerhardt, & Nida-Rümelin, 2008) of consciousness have recently started to appear. Taken together, the literature on the evolution and function/s of consciousness gives the impression that there is much research still to be done, but also that progress can be made and that an increasing number of researchers are working on the topic.

My analysis of this literature identified a multitude of hypotheses about the evolution and function/s of consciousness, often for particular kinds of consciousness, and thus agrees with Dennett's (1991) quotation in this respect. As I have already described the literature's current state of transition, I concentrate here on works published more than five years ago. Many of the identified hypotheses were difficult to locate because they tend to be stated in a few sentences somewhere in a publication. Moreover, individual researchers have only rarely developed their hypotheses further (Mandler, 2002), nor have others in the field done so (Wilcke, 2003). Among the great diversity of suggested functions of consciousness, some reappearing themes, such as integration (Baars, 2002; Morsella, 2005) and behavioural flexibility, have now been recognised (e.g., Delacour, 1995; A. K. Seth, 2009; Van Gulick, 2009; Wilcke, 2003). Even so, the overall picture of the literature is one of many isolated hypotheses for the function/s or evolution of consciousness, which are yet to be explored and, if promising, developed.

The majority of these hypotheses are insufficiently supported, as authors typically offer little or no evidence. Dennett (1991) called evolutionary hypotheses about consciousness *speculations*, as did Welshon (in press) very recently; Morsella (2003) complained that there are hardly any serious hypotheses on the function of phenomenal consciousness. Of course, evolutionary explanations are essentially historical narratives (Dennett, 1991; Mayr, 1983), and hypotheses play an important role in research, but we also need evidence if we want a scientifically respectable evolutionary explanation of consciousness.[1,2] Furthermore, when authors do provide support for their hypotheses, they frequently fail to report the means by which this support was determined (Wilcke, 2003). A significant portion of hypotheses seem to be based exclusively on the inspection of the author's own consciousness, which can be problematic for finding the evolutionary function/s of consciousness, for example, because "Nature does not tell us what our organs are for" (Barlow, 1980, p. 82; see section 4.1). The question therefore is, even with the encouraging recent trend in the literature,

[1]Although evolutionary hypotheses without evidence are sometimes disparagingly labelled *just-so stories*, they are only damaging if they are accepted or dismissed prematurely or if they cannot be investigated scientifically.

[2]Another common use of evolutionary stories about consciousness is reasoning about philosophical theories of its nature (Polger, 2007).

whether the dearth of evidence and the potential difficulties with means of inquiry are due to the nature of consciousness itself.

Before exploring this question, however, two criticisms levelled against all evolutionary explanations of consciousness to date deserve attention. According to Flanagan and Polger (1995), the problem with existing hypotheses is that they suggest functions for consciousness that can be performed unconsciously (by us, by computers, or in principle)—without explaining why conscious organisms nevertheless prevailed over their unconscious conspecifics. Flanagan and Polger demanded a credible explanation either of the success of conscious organisms that competed with zombies or of the absence of competing zombies. Yet for zombies to have existed at the same time, two mechanisms that produce exactly the same behaviours would have had to occur simultaneously in a biological population, which seems highly unlikely. If conscious organisms did coexist with behaviourally identical zombies, natural selection could not have distinguished between them, leaving chance as an explanation for why we are conscious. Researchers have indeed mostly failed to pay attention to competing unconscious conspecifics, but neither should this be the sole determinant of the worth of evolutionary hypotheses nor are zombies the crucial point: If consciousness had a (nonunique) causal role function, its possessors probably differed behaviourally from their less conscious conspecifics.

The second criticism concerns an alleged "fundamental flaw in every proposal for an evolutionary explanation of consciousness" (Polger & Flanagan, 2002, p. 30). Polger and Flanagan (2002) accused all adaptationist hypotheses about consciousness of explaining its evolution by claiming that it is necessary for some ability of the organism. I understand their criticism as being based on the irrelevance of necessity to natural selection, as Williams (1966) explained: "Selection has nothing to do with what is necessary or unnecessary, or what is adequate or inadequate, for continued survival. It deals only with an immediate better-vs.-worse within a system of alternative, and therefore competing, entities" (p. 31). This criticism applies not only to contingent necessity, because the relevant ability could have been realised without consciousness, but also to (biologically unlikely) strict necessity, because the organism could have managed to survive without the ability. It is difficult to establish which authors of adaptationist hypotheses about consciousness committed this fallacy and which merely used the term *necessary* imprecisely. Either way, many of these hypotheses can be corrected by replacing *necessary* with *advantageous* (Williams, 1966), so that consciousness is claimed to be, for example, advantageous for certain kinds of learning instead of necessary.

A third criticism is worth mentioning, even though it does not target all evolutionary explanations of consciousness. Polger (2007) pointed out that many theories of consciousness claim that a specific ability cannot be exercised unconsciously. If so, an evolutionary explanation of consciousness would not be indicated (over and above an evolutionary explanation of the ability for which it is strictly necessary; see section 1.4). Strict necessity may, however, be claimed less often than appears to be the case. On a more charitable and constructive reading, especially of terms such as *necessary* in a nonphilosophical context, many hypotheses simply advance potential advantages of consciousness in us. For example, in suggesting a method for identifying the function of phenomenal consciousness, Morsella (2003) assumed that "phenomenal states can accomplish something that unconscious processes, for some reason, cannot" (p. 439). Yet he explicitly allowed for the possibility of functional-zombie brains made of artificial neutral networks and, hence, could not be claiming strict necessity. Even those hypotheses that do may be modified to claim contingent necessity, or rather advantageousness, of consciousness for the same abilities.

Particular difficulties scrutinised. Giving a scientifically respectable evolutionary explanation of consciousness is very difficult. It requires, for example, knowledge of the evolutionarily relevant units of consciousness and their behavioural effects, along with criteria for consciousness in other organisms. Providing support for plausible evolutionary hypotheses about consciousness is further hampered by a lack of direct fossil records and limited variability within current species, among other things. It might be possible to infer the presence of consciousness in our ancestors indirectly, for instance, from prehistoric skulls and artefacts and from present-day animals (Polger & Flanagan, 1999; Welshon, in press). Similarly, heritability could be implied by congenital deficits, such as congenital deafness or colour blindness (Polger & Flanagan, 2002) if these left unconscious perception intact. Such difficulties with evolutionary explanations and their partial solutions are the same for consciousness as for many other psychological and even biological phenomena (Polger, 2007; Polger & Flanagan, 1999). However, "the task of giving an adaptation explanation for consciousness inherits those difficulties intrinsic to adaptationist explanation, and complicates them with all the philosophical and scientific problems attendant to consciousness" (Polger & Flanagan, 2002, p. 29). Explaining the evolution of consciousness may thus be particularly difficult due to difficulties with the study of consciousness.

Consciousness is undoubtedly perceived as being difficult to study by scientific means: "In much philosophical, psychological and other scientific writing, a scientific investigation of consciousness has been thought to be difficult, if not impossible" (Velmans, 1993, p. 82). A somewhat discouraging example of this perception is the short entry on consciousness in the fourth edition of *A Student's Dictionary of Psychology* (Hayes & Stratton, 2003), which describes consciousness as "extremely difficult to study" (p. 63). But why is consciousness seen as particularly difficult to investigate scientifically? And is it really? The answer to the latter question depends on the context in which the question is being asked (e.g., research question, point in time, academic discipline), the intended meaning of *consciousness*, and the phenomena with which consciousness is being compared. In the following, I consider some of the reasons for the perceived difficulty of the scientific study of consciousness, compared with the study of other psychological and especially mental phenomena where feasible, and check whether the perceived challenges are indeed unique to the study of consciousness.

Some reasons for the perceived difficulty are historical. Consciousness has at times been considered not amenable to scientific investigation (e.g., before psychology became an independent empirical discipline) or plain unscientific (e.g., during behaviourism), at other times of central importance (e.g., during introspectionism). This varied treatment as a potential research topic might itself raise suspicion. For a long time, "it has been extraordinarily difficult to see it [consciousness] as a variable" (Baars, 2003b, p. 3) and therefore to study it scientifically, maybe because we cannot observe its absence from the inside. The publication of thousands of studies that compare more and less conscious processes, however, shows that this obstacle has been largely overcome (Baars, 2003b). Nonetheless, consciousness has kept reappearing as an unsolved mystery in many contexts, suggesting that it is special in some way (Güzeldere, 1995b). Furthermore, a lack of research progress might indicate that we will never understand certain aspects of consciousness, but this is an open question (see section 1.4). Although progress may have been slow and unsteady, none of the historical reasons mean that consciousness is necessarily difficult to investigate today: Even if consciousness *was* very difficult to study, as many scientific phenomena have been, recent conceptual and methodological advances, besides a shift in attitude, could have changed this.

Two further sets of reasons for the perceived difficulty have to do with influences on researchers and their handling of *consciousness*. First, because consciousness seems essential to who we are and how we live, it is likely to bias researchers, for instance, against declaring it useless (see Chapter 3). Relatedly, naive intuitions based on folk psychology influence the assessment of scientific answers about consciousness (Banks, 1993), and familiarity with their own consciousness may influence the questions researchers ask in the first place. The degree of its subjective significance does set consciousness apart from other psychological phenomena, but influences from intuitions and personal familiarity are also common in their study. Second, Banks and Farber (2003) viewed *consciousness*'s referring to a variety of distinct phenomena as one of two "special methodological challenges for scientific investigation" (p. 3). Memory, emotion, and attention divide into subtypes too, though which aspect of consciousness is being studied may be specified less commonly (Delacour, 1995), possibly because the distinctions are not as well-established. However, all of these concerns can be counteracted by researchers and most of them are not unique to consciousness.

Matters of access to consciousness are generally counted among the most serious difficulties for its scientific study (e.g., Banks & Farber, 2003; Chalmers, 2004; Güzeldere, 1995b; Mandler, 1975; Overgaard, 2003; Wright, 2007). One major concern is that researchers cannot directly observe conscious experiences in research participants, to which the standard response is that scientific observation is often indirect (e.g., Baars, 1988; Banks, 1993; Marcel, 1988). Reports on mental processes can be inaccurate (Nisbett & Wilson, 1977), and reports on conscious contents incomplete (e.g., because of linguistic constraints; Schooler & Fiore, 1997), but the same is true of reports on external phenomena (Marcel, 1988; Velmans, 2007). In any case, the validity of results from reports can be increased, for example, by suitable reporting conditions (Ericsson & Simon, 1980) and intersubjective agreement (Velmans, 1993). There is also the suspicion that a scientific explanation of consciousness "is always doomed to leave something essential to consciousness out" (Güzeldere, 1995b, p. 116), which is based on privileged first-person access to consciousness (e.g., Dennett, 1987). Yet we do not know how much we will be able to explain and should not burden scientific explanations with unreasonable expectations: Consciousness from the inside is to be explained, not thereby generated or experienced (G. M. Edelman & Tononi, 2000). Despite the challenges associated with first-person reports, which are employed in many areas of psychology, such reports are a vital source of information about consciousness (see section 4.1).

To sum up, the research literature on the biological evolution and function/s of consciousness is still in its infancy, though first signs of change have recently emerged. The literature consists largely of many diverse but underdeveloped hypotheses: They are in urgent need of both empirical and theoretical support as well as more attention to evolutionary theory. There is no doubt that giving an adequate evolutionary explanation of consciousness is very difficult. However, my brief examination of why consciousness itself is perceived as particularly difficult to study suggested that several of the reasons are not unique to consciousness or no longer apply. Moreover, the examination did not uncover any clearly insurmountable obstacles to the study of consciousness. Yet there are difficulties that make its scientific investigation challenging, and these are mainly methodological in nature (Banks & Farber, 2003; Overgaard, 2003; Velmans, 1993; Wright, 2007). Overall, then, consciousness and its evolution being difficult to study does not mean that we should not try, only that we should take particular care that our means of inquiry are fit for purpose.

2.2 Why and how to evaluate means of inquiry

Whether we regard the choices to be made in the research process as problems or dilemmas, it is advisable to follow McGrath's (1981) "Rule I: Always *face* your methodological problems squarely" (p. 180). Leaving psychological and sociological factors aside, the identified problems in the research literature on the biological evolution and function/s of consciousness are by and large associated with a failure to use suitable means of inquiry, as explained below. I therefore propose to evaluate means of inquiry that have been used (or suggested to be used) in this context, with the aim of determining which of them are most likely to help advance research and which of them should best be left alone. Once it is clear which means are available and what to expect of them, investigating the evolution and function/s of consciousness should become less difficult. In this section, I first argue that an evaluation of means can make a significant contribution to research in this area, give a brief for such an evaluation, and then describe how a literature search turned up the need to construct an evaluation tool for this purpose.

Benefits of evaluating means. Means of inquiry are central to science—as is evaluation. At the most general level, means of inquiry are the strategies, procedures, and tools that researchers use to answer research questions, whether exploratory or confirmatory. Consequently, means of inquiry are indispensable to scientific investigation, which I take to be uncontroversial, though it is rarely brought to mind in everyday research. Note further that evaluation is pervasive in science, as when researchers distinguish between good and bad theories, hypotheses, proposals, research designs, instruments, data, manuscripts, published research, and so on (Scriven, 1972, 1991, 2001). The importance of evaluation in psychology is acknowledged, for example, by the German Psychological Society's (2005) recommendation to include a compulsory course on evaluation in the research methods education of Master's students (similar as in the superseded diploma programmes). Both means of inquiry and evaluation are thus essential components of scientific research.

However, means of inquiry cannot be taken as givens: "Our knowledge about how to conduct inquiry hangs on the same thread from which dangle our best guesses about how the world is" (Laudan, 1987, p. 29). That is, methodological statements are on an equal footing with statements about phenomena under investigation in requiring support (Laudan, 1984, 1987; promoted in psychology by Proctor & Capaldi, 2001). If supported as reliable (i.e., tending to produce accurate results), means of inquiry not only enable us to answer research questions, they play a crucial role in justifying the relevant statements about phenomena (Goldman, 1979, 1986, 1988). In short, means of inquiry justify theories (Laudan, 1984; Nickles, 1987a; Rescher, 1977). If means of inquiry are found to be inadequate, they and their use may be able to be improved, for instance, by a better understanding of appropriate areas of application (e.g., suitable conditions for reports on mental processes; Ericsson & Simon, 1980; appropriate uses of introspection in the evolutionary study of consciousness; see section 4.1). The evaluation in this thesis assesses the support for means of inquiry; and although it does not make explicit recommendations for the development of each evaluated means, it is formative in providing details about means that I hope will be used as a basis for improving them.

The main purpose of this evaluation is to inform the selection of means of inquiry for use in particular research projects; it can furthermore inform the assessment of such choices and their results. It is important to select means of inquiry carefully because they are more or less useful for answering particular research questions. More specifically, means are, as just indicated, better or worse in themselves, more or less suitable, and also more or less efficient (Nickles, 1987b). My reason for initiating the present project was precisely the observation of large differences in the usefulness of means of inquiry that have been employed in or suggested for the study of the biological evolution and function/s of consciousness. Instead of giving sufficiently detailed examples for this observation here, I refer the reader to the following chapters, in which the evaluation of means identifies many of their strengths and weaknesses. With the resulting knowledge of relevant characteristics of means of inquiry, researchers ought to be able to better concentrate their efforts and resources on those means that are best suited for making progress toward a scientifically respectable evolutionary explanation of consciousness.

Assistance with the tricky task of selecting the best means of inquiry should be welcome, as it is often simply not obvious which means of inquiry should best be used to address a particular research problem in consciousness studies. This is not surprising in an emerging multidisciplinary research area that is concerned with a phenomenon whose study is generally accepted as methodologically difficult. Besides means of inquiry that are commonly used in the relevant disciplines, there is an increasing number of means available specifically for studying consciousness, but little methodological consensus (A. K. Seth, Dienes, Cleeremans, Overgaard, & Pessoa, 2008). Evolutionary investigations of consciousness additionally require adequate means for getting at evidence for distant past events, which may be the most foundational issue in evolutionary behavioural sciences generally (Gangestad & Simpson, 2007a). An evaluation of available means of inquiry into the evolution and function/s of consciousness can make researchers' selections and assessments easier and, as already pointed out, better.

What is more, the problems in the research literature discussed in the previous section strongly suggest that an evaluation of means of inquiry is likely to be particularly beneficial in the present context. Poor awareness of the existence of relevant means of inquiry and their characteristics and the ensuing failure to use suitable means are, I suspect, a major source of several of the identified problems, including the use of inappropriate means, the failure to specify employed means, the underdevelopment of hypotheses, and the insufficiency of support. This claim is difficult to prove, yet it is clear that an evaluation of means can help solve each of these problems, namely by examining which means are available for which research question and, more importantly, by determining which means are best, most suitable, most efficient, and hence most promising for contributing substantially to an evolutionary explanation of consciousness. With initial work in the area completed and—going by the increase in research interest—a potentially productive period ahead, such an evaluation of means of inquiry is timely.

Brief for this evaluation. For the reasons given above, I propose to evaluate the promise of means of inquiry into the biological evolution of consciousness in this thesis. As explained in the Introduction, *means of inquiry* is a broad term that encompasses numerous different research procedures and tools for increasing our knowledge about the phenomena under investigation. To enable a consistent evaluation of moderate generality, this evaluation should, based on the foregoing discussion, focus on the following kinds of means of inquiry: research strategies, research methods, and arguments. When selecting such a means of inquiry for a particular study, what is of most interest, apart from situational factors (e.g., expertise, funding, facilities, time frame), is how much the means will likely contribute to the research objective that is directly related to an increase in knowledge. Accordingly, we here want to know how promising a means of inquiry is for helping to construct an evolutionary explanation of consciousness.

This evaluation of the promise of means of inquiry should have the following characteristics: It should be systematic and well-documented so that its procedures are transparent and its results replicable (unfortunately this cannot be taken for granted in evaluation; cf. Scriven, 2001). Moreover, a multicriterial evaluation is preferable to reliance on a single evaluative criterion, not only because the potential contribution of a means presumably depends on more than one factor, but also because an evaluation on several criteria can teach us more about the evaluated means (e.g., how to better match means to particular research situations, how to improve means). In addition to supplying informative descriptions of means of inquiry, this evaluation should be normative (Laudan, 1987; Nickles, 1987a) or, more precisely, prescriptive (Scriven, 1991): It can offer important advice, provided it is comparative (cf. Laudan, 1981), on which one or more means of inquiry should best be used for which research objective. By presenting such an evaluation, this thesis performs the three major tasks of methodology (Nickles, 1987b, 1989; adapted from Reichenbach, 1938), namely describing, criticising, and advising on means of inquiry, within the evolutionary study of consciousness.

To clarify, I am neither promoting the exclusive use of a single means of inquiry nor claiming that we can know in advance which means are best. Researchers need multiple means both to address all aspects of an evolutionary explanation of consciousness and to enable methodological triangulation (D. T. Campbell & Fiske, 1959; Mathison, 1988) when studying one aspect, ideally with the most promising means of inquiry. Naturally, we cannot now determine which means of inquiry are actually best in the present context, that is, which means will turn out to have been the most promising ones (or successful high-risk alternatives). Recommendations of means based on indicators such as past performance and logic can nevertheless advance research, for instance, by leaving less room for the blind continuation of research traditions, personal biases, and oversights and by allowing better informed choices of means of inquiry and assessments of their use.

When deciding how to evaluate means of inquiry, an important issue is the level of generality of the means to be evaluated as well as that of the evaluation itself. As already noted, a moderate level of generality is best suited for the means in this evaluation, as it would be for means of inquiry more generally: Means that are too specific cannot be employed in other research projects without major modifications, yet the most useful means of inquiry are typically content- and context-specific (Nickles, 1986, 1987b). As to the evaluation of promise, it should at least be specific to the promise of means of inquiry; however, to assess the likely contribution of a means, the evaluation needs to take into account the research objective and the maturity of the area of research, including other available means. Pulling in opposite directions also are requirements on the evaluation to be applicable to various means of inquiry and to have clearly communicable results (for easy comparison of the many evaluated means) and demands for details on the strengths and weaknesses of evaluated means (to facilitate a more appropriate use of chosen means and their improvement). It is thus a challenge to pitch the evaluation of promise at an appropriate level of generality and detail, which in turn affects how much help we can expect from the relevant literatures.

The need to construct a tool. Although the evaluation of means of inquiry is a frequently performed and unavoidable routine activity of researchers (e.g., when planning a study or reading research reports), little has been published about pertinent evaluative criteria. While other authors have tended to discuss general ways of evaluating means at most, such as by intuition, logic, or historical evidence (Laudan, 1987; Proctor & Capaldi, 2001; see also Nickles, 1987a), Rescher (1977) advocated effectiveness as *the* evaluative criterion for methods (as did Laudan, 1987) and also mentioned efficiency, rationality, and suitability. Further leads come from the following two lists of evaluative dimensions and criteria: First, Nickles (1987b) suggested that, "since methods govern research processes, methods can be evaluated along three primary dimensions: the suitability of research *goals*, the efficiency of the *processes* in producing research products (e.g., problem solutions), and the quality of these *products*" (p. 42). The efficiency of means of inquiry has in fact been neglected by theorists, but the means themselves should not be judged by the research goal for which they are employed and only by their products' quality if they are applied properly. Second, Nickles (1989) proposed that scientists might evaluate the power of heuristics, depending on the context, by the heuristics' reliability, the difficulty of problems they can solve or avoid, their efficiency, their scope of application, and the incorporation of domain-specific knowledge. These criteria and Rescher's are important considerations, on which I comment when describing the construction of my evaluation tool in the next section. Developing these criteria involves determining how they can be measured and how they relate to the promise of means of inquiry.

Originally I had hoped to find one or more at least occasionally cited lists of method evaluation criteria in the literature, just as there are well-known criteria for theory evaluation. (Even better would have been a well-developed evaluation system, such as the theory of explanatory coherence by Thagard, 1989, 1992.) A first step in increasing my search radius was to check how well theory evaluation criteria might generalise to the evaluation of means of inquiry. Compared with such criteria as predictive accuracy and explanatory power, Kuhn's (1977) standard criteria of theory adequacy, for example, seem reasonably suited for this purpose: Good means should arguably be consistent (especially internally), produce accurate results, and be fruitful; they should simplify the complexity of the data and, as suggested by Nickles (1989), have a broad scope. This shows that Kuhn's five criteria can, with some imagination, be adapted to means of inquiry, after which these criteria, like the ones for means of inquiry, would need to be operationalised and related to promise.

Unlike most philosophers who, in writing about the evaluation of scientific theories, have focused solely on their acceptability (Laudan, 1977; Whitt, 1992), a few have addressed their pursuit worthiness. Among them is Laudan (1977, 1981), who equated a theory's promise with its progressiveness, measured by the difference in problem-solving effectiveness between its most and least (or a less) recent version (where problem-solving effectiveness is the number and importance of solved minus unsolved problems) or by the rate of such progress. In contrast, McMullin (1976) held that promise could be determined by "examining the theory here and now, and estimating its imaginative resources ... for future extension and modification" (p. 400) that will allow the theory, for instance, to resolve remaining problems or incorporate new areas. A theory's promise may thus lie in its potential to improve its performance on standard criteria of theory acceptance (Nolan, 1999), suggesting a corresponding evaluation. Whitt's (1992) much more detailed account of the appraisal of theory promise is too theory-specific to be useful here, but worth mentioning for recognising that the evaluation of promise depends on both historical dimensions (e.g., a theory's conceptual viability and empirical fertility over a period of time) and formal dimensions (e.g., the analogy and experimental strategies of a theory's heuristic). Furthermore, the following factors relevant to means of inquiry can be drawn from Nickles's (1989, 2006) discussions of heuristic appraisal in science: problem-solving capacity, feasibility, usefulness, costs, and risks. Overall, though, philosophers of science have neglected the evaluation of promise (for many reasons; Nickles, 1989, 2006), and the majority of what has been written about the appraisal of theory promise is difficult to apply to means of inquiry.

Perhaps we can shed light on what makes means of inquiry promising to researchers by considering how students are taught when to employ which means. Despite the relatively broad and intensive training of psychology students in research methods (Breakwell, Hammond, Fife-Schaw, & Smith, 2006), their introductory research methods textbooks are, for the most part, silent on research strategies and heuristics for choosing and evaluating research designs (Evans, 2005). They typically recommend matching methods to research questions, describe some of the methods' strengths and weaknesses, and discuss factors such as ethical issues and different types of reliability and validity (e.g., Graziano & Raulin, 2000; Leary, 2004; Rosnow & Rosenthal, 2005). The general concepts of reliability (roughly, consistency or precision) and validity (roughly, effectiveness or accuracy) are highly relevant to the evaluation of such means of inquiry as arguments and evolutionary methods, but much less so the aforementioned traditional types of reliability and validity for measures and experiments (see Lissitz, 2009, for current perspectives on test validity). The textbook advice is in any case not sufficient for choosing means of inquiry, because, for example, many additional factors affect the selection of means (Golden, 1976; Kulka, 1981) and different methods used for addressing the same research question may produce different results (i.e., method variance; D. T. Campbell & Fiske, 1959; Wilson & Lipsey, 2001). For whatever reasons the advice on selecting means is inadequate,[3] it contributes little beyond the general considerations of reliability, validity, and ethics to the evaluation of means in the present context.

Turning now to evaluation advice, guidance on the evaluation of research, in particular its means, is generally along the lines of the topics found in research methods textbooks—whether this guidance is for students evaluating published journal articles (e.g., Dunbar, 2005; Pyrczak, 2008), for researchers peer-reviewing manuscripts (e.g., Joint Task Force of *Academic Medicine* and the GEA–RIME Committee, 2001; Sternberg, 2006), or for anyone determining the methodological quality of studies for inclusion in systematic reviews (e.g., Fink, 2005; Lipsey & Wilson, 2001). Given that the means of inquiry to be evaluated are more diverse than those on which the above advice is given, as again exemplified by arguments and evolutionary methods, it might pay to look further afield for advice. Some engineering design textbooks, for instance, describe comparatively sophisticated and potentially suitable evaluation techniques (e.g., Cross, 1994; Hyman, 1998; Pahl, Beitz, Feldhusen, & Grote, 2007). It is more difficult to bring the literature on evaluation, nowadays largely dedicated to the evaluation of social programmes, to bear on the evaluation of means of inquiry. However, what the evaluation literature does provide as helpful starting points are lists of potentially relevant generic values (e.g., Guba & Lincoln, 1981; Scriven, 2007a). Moreover, works on evaluation, such as Scriven's various writings (e.g., 1991, 2007a, 2007b) and Davidson's (2005) useful manual, can teach us how to conduct an evaluation of means from scratch.

An exception to these general statements about the evaluation literature is the evaluator Coryn's (2007) search for criteria of good research. He examined researchers' own criteria and the norms and standards of research that are embodied in scientific method, in Merton's (1942/1949) ethos of science, and in the ideal of methodological rigour as judged by the validity, reliability, and objectivity of quantitative studies and the trustworthiness, dependability, and confirmability of qualitative studies. Based

[3]A widespread lack of attention to the topic is suggested by the fact that a few books do provide social science students with more detailed information on how to choose means (e.g., Blaxter, Hughes, & Tight, 2006; Evans, 2005; Golden, 1976; Mitchell & Jolley, 2007). Note that most of the additional factors presented concern the preferences of particular researchers, those of other parties (e.g., collaborators, funders, reviewers, peers), and the availability of resources, none of which can be easily evaluated at the more general level intended here (see section 2.3).

on these sources and his definition of research, Coryn proposed the following list of necessary criteria of good research, applicable to a research and development centre, a researcher, many pieces of research together, or possibly a research proposal: originality/novelty, significance/importance, relevance, fecundity, uniformity, validity, replicability, and ethicality. Several aspects of his criteria descriptions are, after differing degrees of adaptation, also necessary properties of good means of inquiry: Means should be fit for purpose (relevance) and ethical (ethicality); they should be logical and produce well-supported inferences (validity); and they should be repeatable (replicability) and perform well consistently (uniformity). Together with resource economy, also mentioned by Coryn, these selected and modified criteria are comparable with Rescher's (1977) and Nickles's (1989) criteria and will likewise be taken up in the next section.

Before moving on it is worth checking whether the tools used in other evaluations of multiple means of inquiry in psychology and related fields can be adapted for the present purpose. The most common kinds of such published evaluations are characterisations of research methods in terms of their particular strengths and weaknesses (e.g., Cooke, 1994; Crothers & Levinson, 2004), simulation studies of the performance of data analysis techniques (e.g., MacKinnon, Lockwood, Hoffman, West, & Sheets, 2002; Velicer & Fava, 1998), and empirical comparisons of typically two to three means of inquiry (e.g., Green, Rafaeli, Bolger, Shrout, & Reis, 2006; Seal, Bogart, & Ehrhardt, 1998). Apart from supplying examples of positive and negative features of means of inquiry, these evaluations are, unfortunately, of no help to a systematic but time-limited evaluation of the promise of many means of inquiry in the under-developed research area of interest here.

A few published evaluations of means of inquiry, however, are a step closer to the evaluation outlined in the brief above. The authors of these evaluations have supplemented their descriptions of particular strengths and weaknesses of several means of inquiry each with a table containing both evaluative criteria and evaluated means, as illustrated in Table 2.1 (for other examples see J. S. Carroll & Johnson, 1990, Table 8.1; Mitchell & Jolley, 2007, Table 7.3). Such systematic and usually more comprehensive overviews enable a direct comparison of all means of inquiry on each criterion. Table 2.1 is also interesting because Klahr and Simon (1999) employed what they considered "the criteria generally used for evaluating research methods in general and psychological research methods in particular" (p. 528). These criteria are, debat-

Table 2.1

Evaluation of Approaches to the Study of Scientific Discovery (Klahr & Simon, 1999)

Evaluative criteria	Historical studies	Laboratory studies: Exploratory	Laboratory studies: Controlled	Direct observation	Computational modelling
Face validity	++	−	−−	++	−
Construct validity	−	−	+	+	+
Temporal span & resolution					
Short & fine-grained	−	++	++	+	++
Long & coarse-grained	++	−−	−−	−	+
New phenomena	++	++	−	++	+
Rigor & precision	−	+	++	+	++
Control & factorability	−−	−	++	−	++
External validity	−	−	+	−	+
Social & motivational factors	++	−−	−−	++	−

Note. ++ = high; + = moderately high; − = moderately low; −− = low. Adapted from "Studies of Scientific Discovery: Complementary Approaches and Convergent Findings," by D. Klahr and H. A. Simon, 1999, *Psychological Bulletin*, *125*, p. 530. Copyright 1999 by the American Psychological Association.

ably, desirable features of means of inquiry in psychology, yet they are, where relevant, less pressing at the current state of research into the evolution of consciousness than the more basic dimensions introduced in the next section. Criteria-by-means tables can in any case be useful for the comparison and possibly even selection of means of inquiry, although they normally make no provision for combining the subevaluations in the cells of the table into an evaluative conclusion.

In summary, research on the biological evolution and function/s of consciousness could profit from a systematic evaluation of the promise of relevant means of inquiry. The resulting information and advice, if taken on board, have the potential to significantly enhance the selection of means of inquiry—as well as their use, assessment, and development—and thereby to help remedy several of the biggest problems apparent in the corresponding literature. However, there is no suitable or easily adaptable evaluation system or set of evaluative criteria available for conducting such an evaluation, probably because the topic has not received enough attention and because the best means of inquiry tend to be content- and context-specific, as is by implication their appropriate evaluation. This is why I needed to construct a tool for evaluating the promise of means of inquiry in the present context.

2.3 Construction of the evaluation tool

Left wondering how to evaluate the promise of means of inquiry into the biological evolution and function/s of consciousness, we should keep in mind the mantra of matching research methods to research questions, as it applies equally to evaluation questions (see, e.g., Owen, 1999; Rossi, Lipsey, & Freeman, 2004). In tailoring a suitable evaluation tool, there were, of course, further aspects that needed to be taken into account, such as the purpose of the evaluation, the evaluation brief (both covered in the previous section), and available resources. Accordingly, this section first sets out the evaluation questions, followed by the requirements for both means of inquiry to be evaluated and evaluative criteria. I then describe the steps that were involved in the construction of the evaluation tool and give an overview of the tool, which is specified in the following section and applied in the remainder of the thesis. At the end of this section, I take a look at the tool's validity, consistency, and objectivity to allow an initial assessment of its trustworthiness.

Evaluation questions and requirements. The primary evaluation question here was which means of inquiry are most promising for increasing our knowledge about the biological evolution of consciousness. A means of inquiry that is more promising than most other means may, however, still be unpromising (i.e., high relative but low absolute promise) and hence not worth pursuing, which is why it was also important to determine *how* promising each means is. The evaluation should additionally provide absolute and relative subevaluation results on every evaluative criterion to facilitate the improvement of means of inquiry as well as a better fit of means to specific research situations (e.g., available resources, risk aversion). Note that because the evaluation was not about a particular researcher or project, it could not easily include factors that differ between researchers and projects if clear evaluative conclusions were to be reached (their inclusion would have considerably complicated the evaluation and probably led to a large number of conditional conclusions). Furthermore, I intended this evaluation of promise to be valid for the potential use of evaluated means of inquiry within a decade (and then to require review), thus giving sufficient time to conduct studies based on the evaluation findings while avoiding unnecessary speculation in a rapidly changing field of inquiry.

To be eligible for this evaluation, means of inquiry had to have been either evidently used or suggested to be used in the investigation of any aspect of the biological evolution of any aspect of consciousness. The means to be evaluated were research strategies, research methods, and arguments, according to the evaluation brief. The exclusion of means of inquiry that had not been used or suggested served to confine the evaluation to means that had been considered promising in at least one research publication. The judgement whether means had been (suggested to be) applied to consciousness should follow the respective authors' claims, as explained in section 1.1. *Any aspect of the biological evolution of consciousness* should be taken broadly to include, for example, its current function/s (as argued in section 2.4), excepting only epiphenomenalism (already discussed in section 1.3; arguments that are more directly relevant to the evolution of consciousness are evaluated in Chapter 3). Besides fulfilling these requirements, means of inquiry needed to be ethical, operationalised as given or expected approval by an institutional ethics committee, to qualify for an evaluation of their promise.

The criteria used in evaluating means of inquiry should be those properties of the means that make them promising in the present context. Among the basic requirements enumerated by Scriven (1994, 2007b) for such a list of evaluative criteria are that the criteria should be clear, nonoverlapping, and at a comparable level of generality. Criteria were to be preferred over indicators, where possible, because a definitional connection with promise is better than merely an empirical one (i.e., a correlation; Scriven, 1991, 2007b). Moreover, the list of criteria should contain neither significant omissions nor superfluous criteria; each evaluative criterion should help to distinguish more and less promising means of inquiry here. The final requirement for evaluative criteria was that the degree of their presence in means of inquiry needed to be determinable by me and by other researchers who are familiar with the topic area and means of inquiry. With this framework in hand, I now describe how I constructed the promise evaluation tool for means of inquiry into the evolution of consciousness.

From values to synthesis. The three main tasks in the construction of the evaluation tool were to identify the relevant values, to set out how the performance of means of inquiry on each criterion is to be measured and rated, and to define the synthesis procedure for combining these subevaluations into an overall conclusion about the promise of a means. A good place to start identifying relevant values was the characterisation of means of inquiry given in the evaluation brief above, not least because "methodological value judgments are *typically* built on definitional or quasi-definitional premises" (Scriven, 1972, p. 240). The brief stated that means of inquiry are procedures and tools that are employed by researchers to attain those research objectives at different levels of generality that are directly related to learning more about the phenomena under investigation. This characterisation highlights that means of inquiry are means to an end, which should therefore be central to their evaluation (Laudan, 1984; Rescher, 1977). Thus, a good means of inquiry is one that attains its research aim.

Based on an analysis of the concept of *a promising means of inquiry*, including an initial literature review (roughly the first half of the subsection The Need to Construct a Tool), I constructed a basic evaluation scheme with the following three dimensions:

- ***Adequacy.*** How does the aim of the means of inquiry relate to the research goal, which is giving a scientifically respectable evolutionary explanation of consciousness?
- ***Validity.*** Does the means of inquiry tend to promote its aim, based on plausible arguments and, where possible, empirical evidence from its past performance?
- ***Strengths and weaknesses.*** What additional strengths and weaknesses does the means of inquiry have (e.g., assumptions, rules, specificity, data availability, required resources)?

This evaluation scheme was piloted on the 10 means of inquiry evaluated in Chapter 3. Its application resulted in qualitative evaluative conclusions (somewhat akin to the descriptions of strengths and weaknesses criticised above), which I found unsatisfactory because they were not sufficiently transparent, systematic, and comparable. For example, it was not clear how levels of validity were determined, whether the same standards were used for all means, how means could be compared on their strengths and weaknesses, and how promise conclusions were derived. In my judgement, the evaluation scheme was in need of substantial revision.

To arrive at a more satisfactory evaluation tool, I pursued two routes, both of which were greatly helped by a more extensive consultation of the literature (summarised in the second half of the subsection The Need to Construct a Tool). On the one

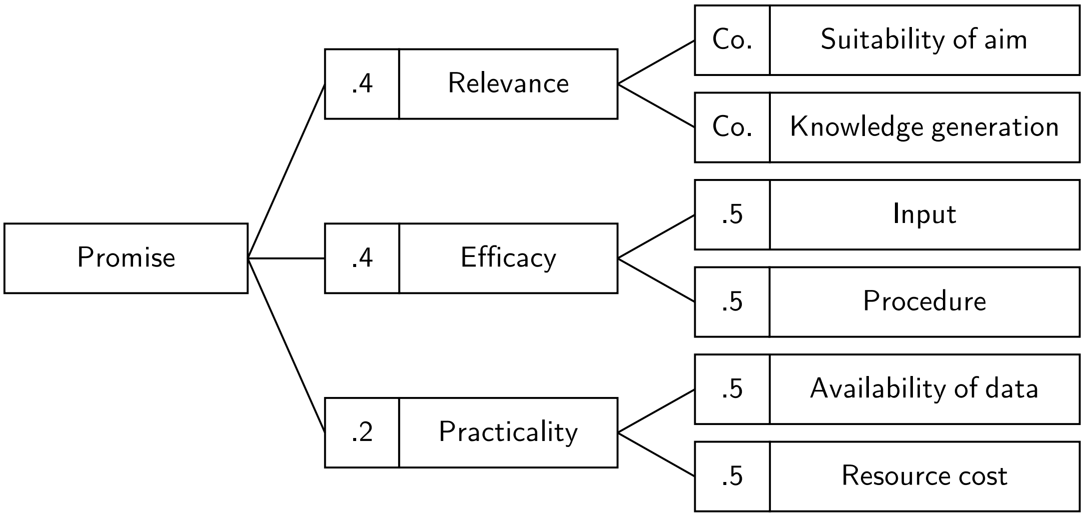

Figure 2.1. Overview of the tool I constructed for evaluating the promise of means of inquiry into the biological evolution of consciousness. Numbers denote weights of evaluative criteria (right) used in determining a means's performance on connected evaluative dimensions (middle) and of dimensions in determining its promise. Co. = scored in combination.

hand, I fleshed out the evaluation scheme by drawing up scoring rubrics for each dimension and devising a numerical synthesis procedure with cutoffs. However, I was not able to fashion a coherent and defensible set of evaluative criteria. On the other hand, I extracted from the consulted literature and from my own notes all those features of means of inquiry that were potentially relevant here. I completely reworked the resulting list of 34 sets of candidate values recursively seven times, under consideration of the aforesaid requirements for evaluative criteria and the present state of the research literature on the evolution and function/s of consciousness. As will become apparent in the next section, the current version of the evaluation tool, outlined in Figure 2.1, turned out to have more or less similar evaluative dimensions and criteria as the original evaluation scheme, while being significantly more elaborate and systematic.

The evaluation tool is made up of only three evaluative dimensions, each assessed on two criteria, yet all of these are *essential* to the promise of means of inquiry into the evolution of consciousness. For several reasons I decided to focus on the needs (i.e., requirements for satisfactory functioning; Davidson, 2005; Scriven, 1991) and not the wants (i.e., desirable features) of promising means of inquiry, thereby accepting a somewhat incomplete evaluation for now. Needs are naturally much more important than wants (e.g., improvability of a means; see Scriven, 1991, on needs assessment). More importantly, needs are likely to be sufficient at the present state of research in this area to distinguish promising from unpromising means of inquiry, as several of the evaluated means, or at least their applications, have rather elementary problems. Furthermore, the descriptions of applied or suggested means of inquiry often lack detail, already making the evaluation of needs difficult. For a first evaluation study of means of inquiry in an emerging and multidisciplinary area of research and with limited resources, it seems most appropriate to use a no-frills tool on a large variety of possibly promising means.

The tool's synthesis procedure, termed *numerical weight and sum* by Scriven (1991), is an application of multi-attribute utility theory: Performance scores are multiplied by the weight (i.e., relative importance) of the corresponding criterion and then summed to obtain dimension and promise scores. This common procedure is suitable for answering absolute and relative evaluation questions (unlike qualitative weight and sum; Davidson, 2005) about many candidates (Scriven, 1991), but it needs to be handled with care because of some of its assumptions (e.g., independence of criteria, interval scale, compensatory criteria; Scriven, 1991) and because of the false sense of certainty easily imparted by the use of numbers (Pahl et al., 2007). To deal with some of these problems, the present evaluation tool has a small number of evaluative criteria, each with a cutoff and equal weighting (the only exception, which is the weight of the practicality dimension, stays within Scriven's, 2007b, recommendations). In addition, the scoring rubrics all range from 0 (*very low*) to 10 (*very high*) and are constructed with equidistance between points in mind. (Both performance assessment on the criteria and their weightings are explained in the next section.) What has not been built into the evaluation tool is a measure against the tendency to overestimate the certainty of numerical results, but several of the other potential problems with this common and useful synthesis procedure have been mitigated or resolved.

Preliminary evaluation of the tool. To get a first impression of the quality of this quantitative evaluation tool, I now consider tentatively its objectivity, consistency, and validity. Objectivity refers to the absence of bias, that is, the absence of systematic errors (Scriven, 1991). Biased results may derive either from the evaluation tool itself or from its application (Scriven, 1994). Because any bias in the tool depends on how well the tool is justified (Scriven, 1991), I have tried my best to construct a well-supported tool, which I have presented to psychology and evaluation audiences and revised in light of their comments. In order to minimise bias in the tool's application, I provide scoring guidelines in the next section, which are meant to render the evaluation more transparent and repeatable. These instructions could be expanded if required, but the risk of their misinterpretation is low in the present evaluation because the author and user of the tool are identical. The tool's objectivity is much higher than that of the original evaluation scheme; it could, however, be further enhanced, for example, through feedback from other researchers and applications of the tool to the same or similar evaluation questions.

An initial indication of the tool's inner consistency (not to be confused with internal consistency in psychometrics) comes from two different assessments of the 10 means of inquiry that were already used in piloting the evaluation scheme (see Chapter 3 for their proper evaluation). I first performed a quick and chiefly global evaluation of the promise of these 10 means with an early version of the current evaluation tool, where neither the evaluative criteria had been finalised nor any weights assigned to them. After the lapse of about six weeks and without knowingly referring to the previous evaluations, I applied a nearly current version of the tool to the same means of inquiry in an analytical pilot evaluation, scoring the means on each criterion and calculating synthesised promise scores. The globally (M = 4.95, SD = 1.09) and analytically (M = 4.92, SD = 0.98) determined promise scores were strongly correlated, $r(8) = .95$, $p < .001$, 95% CI [.80, .99]. Even when taking into account that these pilot evaluations were conducted by the tool's author, the result suggests that the tool's synthesised performance assessments on all criteria measure the same attribute of means of inquiry as a global evaluation of the means' promise.

The next question, then, is that of the tool's validity, namely, whether the evaluated attribute of means of inquiry is in fact their promise. The unanticipated convergence of the original scheme for evaluating the promise of means and the current promise evaluation tool can be taken as a provisional sign of the tool's validity—after all, I constructed the latter anew, as described above, because I was dissatisfied with the former. For want of criteria of promising means in the literature, we can compare the present evaluation tool only to other relevant lists of criteria. The tool's three evaluative dimensions occur in, at times even constitute, some lists of evaluation values (e.g., Foundation for Young Australians & Sharp, 2000; Guba & Lincoln, 1981;

Wells, 1986, as cited in Owen, 1999), though with efficiency instead of practicality. Stronger support comes from Rossi et al.'s (2004) linking of the tool's dimensions to the choice of methods: "The [evaluation] methods selected must be practical as well as capable of providing meaningful answers to the questions with the degree of scientific rigor appropriate to the evaluation circumstances" (p. 33). Moreover, Coryn (2007), Nickles (1989), and Rescher's (1977) values of good means of inquiry or good research (see section 2.2) together speak to the majority of the present tool's evaluative dimensions and criteria, notably relevance (Coryn, Rescher) with aspects of knowledge generation (Nickles), efficacy (Rescher) with procedures (Coryn), and resource cost (as *efficiency*; Coryn, Nickles, Rescher). The main omission is again the availability of data, which is more pertinent to the promise of means of inquiry (or research) than to their quality alone and a particular problem in investigating the evolution of consciousness. Beside this tentative evidence for the tool's essential criteria of promise, the results of its application to means of inquiry in Chapters 3 to 5 as well as analyses of the means' future performance should help to better determine the tool's validity.

To recapitulate, I constructed a tool for evaluating the promise of means of inquiry into the biological evolution of consciousness because I could not find a suitable evaluation system or set of evaluative criteria in the literature. The evaluation tool consists of only those criteria that are essential to the promise of means of inquiry in the present context and uses a numerical weight and sum procedure for synthesising subevaluations. Its aim is to provide both absolute and relative promise and subevaluation results for each evaluated means of inquiry. A first check of the tool's validity, consistency, and objectivity revealed preliminary evidence of its trustworthiness. What we need to know, though, before the evaluation tool can be applied is how to go about assessing means of inquiry on each criterion.

2.4 Assessing promise by 3 x 2 criteria

So how can we determine the promise of means of inquiry practically? The first step in preparing to use the tool just introduced is to identify the means of inquiry to be evaluated. In the present project, this was done by searching the research literature on consciousness for means of inquiry and checking whether they fulfill the requirements stated in the previous section. It is then crucial to pinpoint the aim of each means

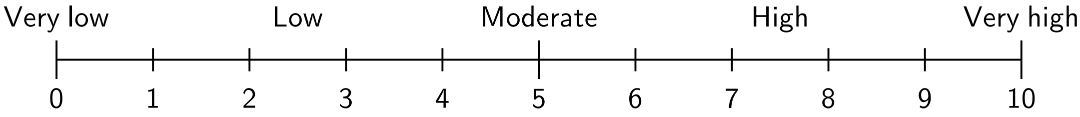

Figure 2.2. Scale used in assessing means of inquiry on evaluative criteria and for interpreting dimension and promise scores. Each label applies to scores up to 1 point higher and lower than its centre.

of inquiry in context, as much of the evaluation relies on correct knowledge of these aims. When the aim of a means of inquiry is not obvious, it can be helpful to try writing out a means-end relation of the form "if one's goal is y, then one ought to do x" (Laudan, 1987, p. 24). Once the means of inquiry and their aims are identified, the tool can be applied as specified after the following introductory remarks.

The guidelines set out below for scoring means of inquiry with respect to the tool's criteria are what Scriven (2007b) called *weakly sequential*: Adhering to their order is for the most part not necessary, but doing so is beneficial. For example, the relevance of means of inquiry is to be assessed first because it has already been judged provisionally for the inclusion decision and sets the scene for the remainder of the evaluation (viz., whether the means are capable of making their relevant contribution and whether this is currently practical). The performance of means of inquiry on each criterion is scored on an 11-point scale ranging from 0 (*very low*) to 10 (*very high*; see Figure 2.2). Because the two criteria in the relevance dimension are scored in combination and required greater precision, I allowed half-integers on this dimension, the same as on the other dimensions. The scoring rubrics describe very low, moderate, and very high performance of means of inquiry and give examples. (Additional descriptions of intermediate performances will be more appropriate if the evaluation tool proves useful in the present project.) Let us now see how the promise of means of inquiry can be assessed on the three dimensions: relevance, efficacy, and practicality.

Relevance dimension. The more a means of inquiry may contribute to the research goal—a scientifically respectable evolutionary explanation of consciousness—the more relevant it is in the present context. What we can realistically expect to achieve is, of course, far from the ideal research outcome, which would be a complete scientific explanation of the biological evolution of all aspects of consciousness in all organisms. The importance of the relevance dimension is illustrated by the editorial comment that "among the weaknesses most often cited by reviewers of manuscripts ... is the mismatch between the research question and the research methods and analysis techniques employed to answer the question" (Lester, 1996, p. 130), and this despite what I referred to as *the mantra of matching research methods to research questions*. The assessment of the relevance of a means of inquiry can be broken into two steps: First one classifies the means according to how suitable its aim makes it for answering the research question of how consciousness evolved. Because means with the same general aim may nevertheless differ in their potential contribution, one then determines the size of the means's contribution in the respective category under the temporary assumption that the means actually works.

By distinguishing five approaches to constructing an evolutionary explanation of consciousness in the literature, or rather by identifying five subgoals thereof, I created a simple classification system for the aims of means of inquiry:

1. Showing that consciousness has a function (independent of what this function may be; *a function*).

2. Identifying current function/s of consciousness (*what function/s*).

3. Showing that consciousness is (or was) an evolutionary adaptation (*an adaptation*).

4. Identifying evolutionary function/s of consciousness (i.e., its survival value/s or adaptive significance/s; *what adaptation/s*).

5. Identifying the evolutionary origin of consciousness and its development since (*origin and history*).

These subgoals include means of inquiry with aims either matching or opposite to the above statements. An example of the latter are deflationary evolutionary explanations, which aim to show that some kinds of consciousness are nonadaptations (see section 5.1). However, opposite aims are rare among means of inquiry evaluated in this thesis, which is why I concentrate on matching aims in the present chapter.

The approximate relevance of the subgoals to the overall research goal was determined relatively, by ranking them based on (a) a specification being better than an existential quantification (i.e., $\exists$, there exists; here, *identifying* approaches are more relevant than *showing-that* approaches of equivalent content) and (b) more evolutionary information being better than less:

$$\begin{array}{ccccc} \text{A function} & <_{\mathrm{a}} & \text{What function/s} & & \\ \wedge_{\mathrm{b}} & & \wedge_{\mathrm{b}} & & \\ \text{An adaptation} & <_{\mathrm{a,b}} & \text{What adaptation/s} & <_{\mathrm{b}} & \text{Origin and history} \end{array}$$

This framework of subgoals for assessing the relevance of means of inquiry is further developed below. First note that although the subgoals can be pursued independently, they are admittedly interrelated, with research outcomes in one approach likely to affect other approaches (e.g., identification of an evolutionary function would probably support the corresponding is-an-adaptation hypothesis). A means of inquiry may therefore appear to have more than one aim, in which case it should be assigned to the eligible category with the highest relevance, unless there is a clear indication that this would lower its promise score. Note also that when considering single instances of an author's use or suggestion of a means of inquiry, the author's goal may not coincide with the present research goal.

Finally, the inclusion of nonevolutionary subgoals in the framework may be questioned, given that the overall research goal is an evolutionary explanation of consciousness. The first nonevolutionary subgoal, showing that consciousness has a function, targets epiphenomenalism. Means of inquiry in this lowest ranked category are not evaluated in the present project, as already indicated in the requirements for means, in order to focus on the many similar and much more relevant kinds of means that are. Conversely, means of inquiry within the other nonevolutionary approach, identifying current function/s of consciousness, could not be ignored. For one thing, the identified function/s may well have been shaped by biological evolution. What is more, many potentially promising means of inquiry cannot be used on consciousness in the distant past but only on the consciousness of living creatures, typically humans. These means of inquiry can provide us with much-needed knowledge about—at least current—consciousness, which may constrain possible evolutionary explanations and suggest hypotheses about how and why consciousness could have evolved.

The relevance of a means of inquiry within a category (also) depends on the size of the means's potential contribution to the research goal, that is, the contribution the means can make if it works properly and is successfully applied. Factors that influence how much and how well-supported knowledge a means of inquiry can be expected to

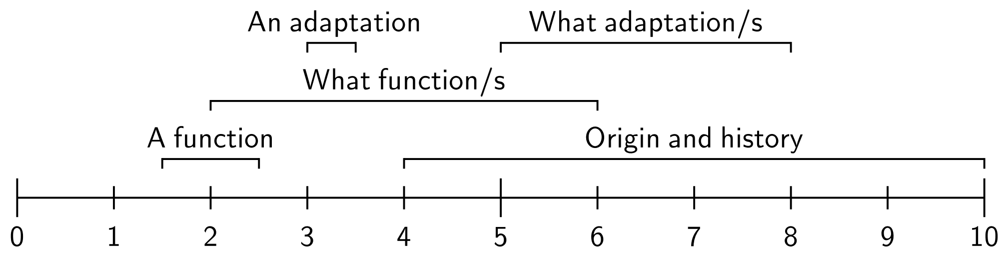

Figure 2.3. Relevance ranges of the five approaches to constructing an evolutionary explanation of consciousness (derived from Table 2.2).

generate include the means's scope of application (e.g., all aspects of consciousness but only in humans) and its specificity (e.g., consciousness not confounded with mind, perception, or attention). The means of inquiry may have been used in or suggested for either the generation or the development of hypotheses, both of which are generally indispensable to research progress. However, because the literature on the evolution and function/s of consciousness largely consists of many speculations with insufficient support, means of inquiry that contribute evidence tend to be particularly important in the present context. Estimating the knowledge generation potential of means of inquiry is difficult, even after having allocated their aims to subgoals, and calls for more concrete advice or at least examples.

To create guidelines for assessing the relevance of means of inquiry, I fleshed out the framework of relevance-ranked subgoals by developing descriptions of what constitutes a small, moderate, and large contribution to the research goal within each category and by assigning numerical relevance values to these descriptions (see Table 2.2). The resulting range of contribution-dependent relevance scores for each subgoal is depicted in Figure 2.3. Apart from striving for a consistent scoring rubric with equidistant points, two further possibilities needed to be considered. Some means of inquiry are likely to lead to relevant knowledge advances beyond their respective aims, for which they receive a *fertility bonus*. For example, the aim of arguments from complex design is to show that consciousness is an adaptation, but they may give additional clues about the evolutionary function of consciousness (see Chapter 3). Other means of inquiry are not sufficiently relevant to the present research goal to be worth pursuing at this time (viz., means with a relevance score ≤ 1.5) and are therefore not included in this evaluation. The relevance scoring rubric provided in Table 2.2 takes into account the most relevant aim of a means of inquiry, the size of the relevant knowledge it may generate within its approach to the research goal, its positive side-effects on knowledge associated with other approaches, and the minimum acceptable relevance.

Efficacy dimension. The more a means of inquiry is capable of achieving its aim, the more efficacious it is. Assessing the efficacy of a means of inquiry thus means assessing how well the means can be expected to work under ideal circumstances. Ideal circumstances are assumed here because the efficacy of a means of inquiry is about the means itself, that is, its capability, and not about the characteristics of particular researchers, the resources they have available, or the data they collect. Factors affecting the application of means of inquiry are better evaluated as part of the means' practicality assessment. In the present evaluation tool, efficacy is assessed on two equally weighted criteria:

$$\text{Efficacy} = .5 \text{ Input} + .5 \text{ Procedure}$$

During the following explanation of these criteria—and in fact during each application of the tool in the remaining chapters—it is important to keep in mind that the tool is for evaluating the means of inquiry themselves, not specific instances of their use. The latter nonetheless inform the evaluation in several ways, especially in the efficacy dimension, as will become apparent shortly.

The most common types of input to means of inquiry are the data to be processed and the premises of arguments. Such input, as far as it is specified by the means of inquiry, is to be evaluated for its quality or evidential strength. For example, data from introspection is strong because it derives from more or less direct observation of conscious experience but also weak because it tends to be biased (for the use of introspection in evolutionary studies, see section 4.1). Although the criteria for consciousness that are used to determine relevant input could be considered part of the input to evaluated means of inquiry, they should only be taken into account in the evaluation if they are specified by the means and differ significantly in quality from commonly used criteria. This is because the focus of the evaluation is on evolutionary means of inquiry, all of which have to rely on criteria for consciousness but generally leave their selection to the user. The evaluation of input should be based on analyses of the means of inquiry and, where available and relevant, records of their past applications. If the strength of the input to a means of inquiry is too low (i.e., input

Table 2.2

Rubric for Assessing the Relevance of Means of Inquiry Within One of Five Research Subgoals and With Potential Additional Fertility

Score	Description
	A function
1.5	Consciousness has physical effects.
2	Consciousness causes physical events that may be relevant to the functioning of some animals.
2.5	Consciousness has a causal role function in certain animal species.
	What function/s
2	Minor current function of consciousness that seems unlikely to be evolutionarily relevant, without any additional details.
4	Current function/s of consciousness with some additional details.
6	Current function/s of consciousness that seem crucial to the functioning of at least some species and may be evolutionarily relevant, with contextual information about its role/s in the organisms' functioning.
	An adaptation
3	Consciousness has (or had) survival value.
3.5	Consciousness either is or has been an evolutionary adaptation in certain animal species because of certain specified features of consciousness.
	What adaptation/s
5	Past or current survival value of consciousness without any additional details.
6.5	Past or current evolutionary function/s of consciousness with some contextual information.
8	Evolutionary function/s of consciousness, covering both past and present, with contextual information.
	Origin and history
4	Partial evolutionary history of consciousness without any contributing factors or other details; for example, a partial description of one stage or development during the origin of consciousness or its more recent history in one species without any causes or effects.
7	Either evolutionary history of consciousness without additional details or partial evolutionary history of consciousness with some contributing factors and other details; for example, the origin of consciousness or a later transition with some reasons for its development.

Table 2.2 (*continued*)

Score	Description
10	Evolutionary history of consciousness with as many details as we can currently hope for, including the origin and following development of all aspects of consciousness until today with contributing factors.
	Fertility bonus
+0.5	Some additional benefits for other approaches to the research goal, such as a plausible new hypothesis, a new approach to a problem, or other leads for future research outside the aim category of the means of inquiry.
+1	Significant additional benefits for other approaches to the research goal, such as evidence with clear and strong implications for other categories.

Note. The bar (i.e., minimum performance for potentially promising means of inquiry) on each criterion is 2.

score $\leq$ 1), the means cannot be promising—in line with the phrase "garbage in, garbage out"—independent of the means's performance on other criteria and dimensions. Table 2.3 shows the evaluation tool's rubric for scoring the strength of input to means of inquiry.

The efficacy of a means of inquiry also hinges on how the means arrives at results or conclusions or, in other words, how it processes and transforms the input into output. Consequently, the quality of the procedure of a means of inquiry needs to be assessed. This quality is positively influenced by, for example, a means making sound assumptions, cogently connecting input to output, and excluding alternative conclusions (except where this is linked to the specificity of the means, which is assessed in the relevance dimension). The evaluation of procedures should, like the evaluation of input, be based on analyses of the means of inquiry and, if appropriate, previous applications. Unsuccessful applications may draw attention to procedural problems, whereas successful applications can be evidence that a means of inquiry works (Nickles, 1986), as long as either outcome is not accidental (Rescher, 1977) or otherwise unrelated to the means itself. A means of inquiry is unpromising when its conclusion does not follow or only very poorly follows from the input (i.e., procedure score $\leq$ 1). The anchors given in Table 2.3 indicate how means of inquiry are to be scored on the procedure criterion.

Table 2.3
Rubric for Assessing the Efficacy of Means of Inquiry on Two Criteria

Score	Description
	Input
0	The type or source of data that the means of inquiry is designed to take as input is very poor, or there is very strong evidence against an essential premise of the means.
5	The input to the means of inquiry allows it to work moderately under ideal circumstances. For example, the data may be somewhat biased or the evidence for some of the premises weak.
10	The quality of the data to be processed by the means of inquiry is very high, or the evidence for all premises of the means is very strong.
	Procedure
0	The procedure of the means of inquiry is very poor due to, for example, a faulty assumption, the irrelevance of the input to the conclusion, or a fatal logical error.
5	The means's procedure tends to lead to the conclusion, but, for example, alternative interpretations are possible or an assumption is questionable.
10	The means's assumptions or implicit premises are sound, and its conclusions follow cogently and exclusively from the input.

Note. The bar (i.e., minimum performance for potentially promising means of inquiry) on each criterion is 2.

Practicality dimension. The more a means of inquiry can realistically be put to use, the more practical it is. Because the present project necessarily abstracts from particular situations, the practicality dimension is limited to the currently available prerequisite data for the use of a means of inquiry and its resource requirements:

$$\text{Practicality} = .5\ \text{Availability of data} + .5\ \text{Resource cost}$$

Note that, in a given research situation, the project's recommendations are straightforward to integrate with pertinent situational factors (unless, e.g., the definition of *promising* as *likely to contribute substantially to the research goal* is rejected as too low-risk). The evaluation question in this dimension, in short, is how practicable the application of a means of inquiry is, irrespective of its relevance and efficacy.

Both the quantity and quality of available data or evidence about consciousness that could serve as input to a means of inquiry help determine whether and how successfully the means can be employed. For example, when evaluating the practicality of the comparative method, we need to estimate how much and how good our knowledge of animal consciousness in different species is or could be soon (see Chapter 5). Information on the input to a means of inquiry from the efficacy dimension is thus useful in assessing the availability of data. The data or evidence do not have to be ready for use now; it is sufficient for the purpose of this evaluation if they are likely to be accessible within a decade. Evidence for data availability normally comes from the research literature, including past applications of the means of inquiry in the present context. A means is currently unpromising if hardly any or only very poor data or evidence about consciousness are available for its application to consciousness (i.e., data availability score ≤ 1). Table 2.4 provides guidelines for scoring the availability of data or evidence as input to particular means of inquiry.

The evaluation tool's final essential criterion is resource cost, that is, the cost of the resources required to employ a means of inquiry in a research project. The resource cost of a means and the contribution it is capable of making determine its efficiency, the only value agreed upon by Coryn (2007), Nickles (1987b, 1989), and Rescher (1977). As this tool's criteria are more basic, the efficiency of any evaluated means of inquiry can be calculated from its relevance, efficacy, and resource cost: efficiency = efficacy/cost; efficiency for the present purpose = $0.5 \cdot$ (relevance + efficacy)/cost. The resource cost of a means of inquiry should be estimated on the scale in Figure 2.2 in terms of time and money, which are the most elementary general resources—others tend to reduce to them, or to depend on the particular research situation, or both (e.g., space, equipment, staff, expertise, effort, reputation, opportunity; Scriven, 2007a). A means of inquiry that requires too much time or money is unpromising (i.e., resource cost score ≤ 1). Table 2.4 assists with scoring the resource cost of means of inquiry.

Once the two criteria in each of the evaluation tool's dimensions have been assessed and the results combined into dimension scores, the final promise score can be calculated as follows:

$$\text{Promise} = .4\ \text{Relevance} + .4\ \text{Efficacy} + .2\ \text{Practicality}$$

The weights subordinate practicality to relevance and efficacy because practicality is

Table 2.4
Rubric for Assessing the Practicality of Means of Inquiry on Two Criteria

Score	Description
	Availability of data
0	The data or evidence about consciousness that is required as input to the means of inquiry is either very unlikely to be accessible within a decade or useless because of its very poor quality.
5	The availability of data or evidence about consciousness enables the means of inquiry to be applied to consciousness, yet the means's contribution clearly stays below its potential. There is (or will likely be soon) either a moderate amount of data or evidence of moderate quality available as input to the means of inquiry, or there is some data or evidence of high quality, or there is much data or evidence of low quality.
10	There is (or will likely be soon) ample good data or evidence about consciousness that could serve as input to the means of inquiry.
	Resource cost[a]
0	Employing the means of inquiry would require excessive time, money, or both. That is, even if its use brought significant benefits, the means is very unlikely to be used because the cost of the required resources would be too large for most funding agencies and research groups.
5	The resources required to employ the means of inquiry are moderate overall: Either much time and little money, little time and much money, or moderate amounts of both are needed.
10	Very little time and money is needed to employ the means of inquiry, because its input consists entirely of readily available data, or all of its premises are well-established, and its procedure is easy and fast to execute.

Note. The bar (i.e., minimum performance for potentially promising means of inquiry) on each criterion is 2.
[a]Resource cost is scored inversely (i.e., a resource-expensive means of inquiry receives a low score) for consistency with all other subevaluations, in which a higher score denotes more of a promise-enhancing feature.

less essential to the promise of means of inquiry, as long as the means' performance is above the minimum acceptable level on the corresponding criteria, and because these criteria differ more across situations (e.g., resource cost may be a lesser concern in an exploratory phase of research or in a very well-funded project). The subordination additionally prevents errors in the less accurate assessment of practicality from unduly influencing the overall promise score. Weights may be adjusted when selecting means of inquiry according to the particular research situation, provided there are compelling reasons to do so. Dimension and promise scores can then be translated into descriptive labels using the scale in Figure 2.2. The results of the subevaluations and the final synthesis should be presented in tables and graphs to facilitate the systematic comparison of means of inquiry on criteria, dimensions, and promise.

Conclusion

This chapter explained the rationale for the research reported in this thesis and described the construction of a tool that permitted the research to be carried out. First I argued that the literature on the biological evolution and function/s of consciousness, which is largely made up of various underdeveloped hypotheses, could be advanced considerably by the use of suitable means of inquiry. To help with the challenge of selecting suitable means of inquiry in this emerging multidisciplinary and methodologically difficult research area, I set out to evaluate the promise of means of inquiry into the evolution of consciousness. Because my literature search for an applicable or adaptable evaluation tool was unsuccessful, I constructed one for this purpose. The tool is underpinned by an evaluative theory (Scriven, 2007b) about what essential properties make a means of inquiry promising in the present context. In brief, a means needs to be relevant, efficacious, and practical, as assessed on two criteria each, using scoring rubrics. Applying a numerical synthesis procedure results in a promise score, which indicates what means of inquiry are most worth pursuing because they are likely to contribute substantially to the research goal: a scientifically respectable evolutionary explanation of consciousness.

An issue that came up repeatedly in this chapter is the evaluation tool's degree of elaborateness. For example, I only constructed the tool because the earlier evaluation scheme was not elaborate enough, but then I focused on essential criteria and left the description of intermediate performances in the scoring rubrics for when the tool has shown itself useful. If readers accept that an evaluation of means of inquiry into the evolution of consciousness should be beneficial, the more likely—and in any case the more damaging—charge is insufficient rather than excessive thoroughness and precision. In reply I emphasise that appropriateness is crucial: Making the evaluation much more detailed and its estimates much more precise would require an unreasonable amount of resources for a first evaluation study of the extent of elaboration and in the area outlined above. We do not need evaluation results with a high degree of certainty at this point; an approximate indication of which means of inquiry are currently most promising, and why, already has the potential to significantly improve research and, hence, ultimately our evolutionary understanding of consciousness.

We are now ready to see the evaluation tool in action. In the following three chapters, I apply the tool to many different means of inquiry, all of which have been used or suggested to be used in the investigation of the biological evolution or function/s of consciousness. To avoid varying all evaluative dimensions at once, I start by evaluating 10 means of inquiry with the same research subgoal, namely showing that consciousness is (or was) an evolutionary adaptation. This subgoal is particularly suited for the first round because it is the best-defined subgoal with the smallest range of relevance scores. In addition, the 10 means of inquiry were the ones used for piloting the evaluation scheme and for comparing global and analytical evaluations with earlier versions of the present evaluation tool. The following chapter thus begins to show what we can learn about the means of inquiry, their promise, and the evaluation tool through its application.

3 Methodological adaptationism without hypothesised functions

What is arguably the most fundamental question about the biological evolution of consciousness is whether consciousness has evolved for a purpose, that is, whether consciousness is (or was) an evolutionary adaptation. All means of inquiry evaluated in this chapter are aimed at providing support for the claim that consciousness is a current (or past) adaptation. Sufficient support for this claim would strengthen the many accounts of consciousness and its function/s that are based on the claim (e.g., Baars, 1988; Crick & Koch, 1995; Koch, 2004). Such support could also justify the search for current causal role function/s of consciousness: Several philosophers have argued for the claim that consciousness is an adaptation, and hence causally affects the fitness of conscious organisms, in order to discredit epiphenomenalism (e.g., Flanagan, 1992; James, 1879, 1890/1907; Nichols & Grantham, 2000). Moreover, empirical evidence and reasoning about the adaptation status of consciousness can yield valuable information for pursuing the other three research subgoals (viz., identifying current and evolutionary function/s of consciousness as well as its evolutionary development), for example, by highlighting the most relevant features of consciousness, establishing constraints on its possible function/s, and suggesting not only that it is (or is not) an adaptation but why.

Approaching consciousness as a potential product of natural selection before trying to exclude nonselective explanations is an instance of *methodological adaptationism.* Although it may turn out that a particular trait was not directly selected for during evolution, adaptationism as a research strategy recommends seeking an adaptation explanation of the trait to begin with (Godfrey-Smith, 2001; Holcomb III, 2001). Such an approach is motivated not only by the successful history of methodological adaptationism—properly applied—since well before Darwin's (1859) *Origin of Species* (Godfrey-Smith, 2001; Mayr, 1983; West-Eberhard, 1992) but also by its current prevalence and increased sophistication in evolutionary biology (Alcock, 2005; Freeman & Herron, 2007; Griffiths, 1996; Zuk, 2002). Another reason for pursuing methodological adaptationism is that adaptation hypotheses are often less difficult to confirm or disconfirm than hypotheses about the operation of nonselective factors (Lewontin, 1978; Mayr, 1983; Godfrey-Smith, 2001). Adaptationist methods can help to identify such factors; for instance, an optimisation model's unconfirmed prediction may point to an overlooked genetic or developmental constraint (Maynard Smith, 1978; Sterelny & Griffiths, 1999). Knowledge of adaptations is typically also necessary for supporting alternative hypotheses, for example, about by-products or vestiges (Buss et al., 1998; West-Eberhard, 1992) and for allowing comparative evaluations of selective and nonselective explanations (Andrews, Gangestad, & Matthews, 2002). Methodological adaptationism is thus an important approach to the evolutionary study of biological traits generally.

This chapter focuses specifically on the evaluation of those means of inquiry that aim to show that consciousness is an adaptation independent of the function/s it may have, hence the title *methodological adaptationism without hypothesised functions*. I first consider the basic evolutionary reasoning that appears to commonly underlie the claim that consciousness is an adaptation. The two arguments evaluated next are characterised by their reliance either on users' own experiences or expert opinion on the adaptation status of consciousness. I then examine five further single-sentence adaptation arguments. The final two evaluations are of comparative arguments, which aim to infer the adaptation status of consciousness from its distribution among today's animals, and arguments from the complexity of consciousness as a feature of special design. The chapter concludes with a comparison of the evaluated means of inquiry based on a table of all subevaluations and a graph of dimensional and promise results.

3.1 Analysis of the argument from evolution

Perhaps it is obvious that consciousness is an evolutionary adaptation: "Consciousness ... is no doubt naturally selected" (Penrose, 1987, p. 266). After all, the most developed cognitive theory of consciousness, global workspace theory, simply assumes that, "like any other biological adaptation, consciousness is *functional*" (Baars, 1988, p. 347); and prominent neuropsychologist Richard Gregory (1998) asked, "why should consciousness have evolved if it is useless?" (p. 1693). In this section I examine a common basis of the presumption that consciousness is a current adaptation by writing out the underlying deductive argument, which I call *the argument from evolution*. A detailed analysis of the argument's premises enables a concise evaluation of its promise at the end of the section.

Gregory's (1998) question provides a clue to the reasoning that frequently seems to be behind the claim that consciousness is an adaptation. The point he presumably wanted to make with his question relies on the following modus tollens, a valid form of argument that denies the consequent of the conditional premise:

$$\begin{array}{l} \text{If consciousness were useless, then it would not have evolved.} \\ \text{Consciousness has evolved.} \\ \hline \text{Therefore, consciousness is useful.} \end{array}$$

By using transposition and adjusting the conditional, the conditional premise can be rewritten as "If consciousness has evolved, then it is useful." Because Gregory employed the term *useful* in the context of evolution, it may be replaced by the term *adaptive*. By substituting the latter in turn with the term *an adaptation*,[1] the general argument for consciousness being an adaptation can be expressed as the following modus ponens, a valid form of argument that affirms the antecedent of the conditional premise:

$$\begin{array}{l} \text{If consciousness has evolved, then it is an adaptation.} \\ \text{Consciousness has evolved.} \\ \hline \text{Therefore, consciousness is an adaptation.} \end{array} \tag{A}$$

The form of this deductive argument is valid, but the argument is only sound if both premises are true. The truth of the conclusion that consciousness is an adaptation therefore depends on the truth of the conditional premise and the truth of the second premise.

Second premise. Taking the second premise first, the following breakdown of the claim that consciousness has evolved is helpful in assessing its truth:

[1] This substitution is only justified when taking into account the antecedent, that is, when assuming that the trait in question has evolved. Although not all traits that have evolved and are adaptive are adaptations (it depends on whether their existence is due to positive natural selection over many generations), using evolution to infer adaptiveness implies a concern with adaptation.

$$\frac{\text{We are conscious.} \qquad \text{We have evolved.}}{\text{Therefore, consciousness has evolved.}} \tag{P2}$$

The argument is intuitively phrased, but its conclusion is at risk because the different contexts of the term *we* in the premises suggest different meanings of the term. When we evaluate whether we are conscious we tend to extrapolate from introspection, concentrating on the inner view of our mental processes. However, when we consider the claim that we are a product of biological evolution we are more likely to think of our outward appearance, a salient feature of our phenotype. The tempting change of meaning from mind to body is a reminder to use terms consistently: The meaning assigned to the term *we* in the first premise must be the same as, or suitably related to, the meaning assigned to it in the second premise. Authors who argue from the existence of consciousness to it being adaptive or an adaptation typically speak of *Homo sapiens* with a first-person plural personal pronoun or possessive adjective in proximity (i.e., *we*, *us*, *our*; e.g., Flanagan & Polger, 1995; Gray, 1971). Hence, we should take both uses of the term *we* in argument (P2) as referring to whole human beings without inattentively sliding in meaning from mind to body.

Both premises of the argument for consciousness having evolved are amply supported from a psychological and a biological point of view, respectively. Although philosophical objections can be raised, the first premise that we are conscious is almost certainly true (which is why consciousness realism and other conscious minds could be assumed in section 1.1). Consciousness exists not only in the mental worlds of individual human beings, it is generated by their brains in interaction with their environments and becomes apparent in their behaviours, such as their descriptions of their own conscious experiences. Regarding the second premise, Darwin (1859) already intended his theory of evolution by natural selection to apply to humans including "each mental power and capacity" (p. 488). Today there is no scientific doubt that human bodies—along with the minds their brains generate—are a product of biological evolution (see section 1.2). Considered separately, both premises are highly probable for humans.

The soundness of argument (P2) now only depends on whether the conclusion follows necessarily from the premises. To clarify the connection between the two premises, a reexamination of the meaning of *we* in each premise is required. Even though we are indeed conscious, consciousness is only part of human beings, which is obvious when it is contrasted with, for example, unconscious mental processes or the rest of the human body that is not directly involved in generating consciousness. Modifying the first premise accordingly, reversing the order of the premises, and taking into account the first examination of *we* gives a more precise version of argument (P2):

> Whole humans are a product of biological evolution.
> Human consciousness is a component of whole humans.
> ___
> Therefore, human consciousness is a product of biological evolution.

Note that by calling consciousness a *component* I do not mean to prejudge whether it is a state, a process, or an activity, only to point out that it does not amount to a complete living and acting organism.

The meaning of the evolutionary premise that we or, in the most recent version of the argument, whole humans have evolved becomes clearer if formalised further into the major premise of a categorical syllogism: "All components of humans are products of biological evolution." Only such a strong version of this premise makes argument (P2) valid without additional premises. To count as a product of biological evolution, a trait or organism must have been shaped by evolutionary processes, either directly or indirectly. However, even an evolved trait or organism may have features that are not products of biological evolution, for instance, if they have been determined by constraints or are too recent for evolutionary processes to have effected significant change. Examples of the latter are such widespread activities of modern humans as reading, typing, smoking, and playing soccer. The strong version of the evolutionary premise is therefore false—not all components of humans have evolved—and the deductive argument (P2) unsound.

The conclusion of argument (P2) may nevertheless be true. To argue that consciousness is probably a product of biological evolution, we could use an inductive version of the argument and supply additional premises with support. Consciousness is more likely to have evolved, for example, if there is evidence that it is not purely a recent cultural product (see section 1.4), that it is intimately associated with the brain, that all humans seem to develop it early in their lives and spontaneously, and that prolonged unconsciousness threatens their survival and reproduction. (Another strategy, namely showing that consciousness is an adaptation, which entails that it has evolved, is not available here because argument [A] reasons in the opposite direction: from evolution to adaptation.) Such an inductive version of argument (P2) would, I believe, warrant the conclusion that consciousness has very likely evolved—otherwise I would not have undertaken an evaluation of means of inquiry into the biological evolution of consciousness.

First premise. Turning now to the conditional premise of the argument from evolution, illustrations of the underlying reasoning show that the emphasis is on what the evolution of consciousness contributes to its adaptation status. After having described how Darwin's ideas influenced the study of consciousness, George A. G. A. Miller (1962/1966) explained that "the very fact that an animal had evolved through thousands of generations implied that the changes were adaptive" (p. 151). Similarly, Popper (1977) asserted that "the Darwinian view must be this: consciousness ... [is] to be regarded (and, if possible, to be explained) as the product of evolution by natural selection" (p. 183). Reasoning like this can also be found in more recent publications. For instance, "if we ... follow the general ideas of evolution, then we would expect to find that some of the effects of consciousness have survival advantages" (Squires, 1990, p. 36). Bringsjord, Noel, and Ferrucci (2002) answered the question whether consciousness is "a big deal" for evolution, presumably meaning whether consciousness is an adaptation, by stating "apparently; after all, we evolved" (p. 111). Likewise reminiscent of Gregory's (1998) rhetorical question discussed in the introduction to this section, Humphrey (2002) concluded that consciousness is biologically useful for the reason that "whatever exists as a consequence of evolution must have a function" (p. 68) because it could not have evolved otherwise. These examples imply that we can infer that consciousness is an adaptation from it having evolved.

To analyse the conditional premise of argument (A), the premise can be usefully translated into a categorical proposition. The original statement "If consciousness has evolved, then it is an adaptation" reads in categorical form "All products of biological evolution that are identical to human consciousness are adaptations." This statement is undoubtedly true if all products of evolution are adaptations. However, we know that some products of biological evolution are not adaptations. Well-known examples of human nonadaptations include the redness of blood (Gould & Lewontin, 1979; Wallace, 1889), the whiteness of bones (Tooby & Cosmides, 1992; Wallace, 1889), heart sounds (Canfield, 1964; Hempel, 1959), the chin (Gould, 1977; Lewontin, 1978), male nipples (Symons, 1979), and the belly button (Buss et al., 1998). Because the general statement that all evolutionary products are adaptations is contradicted by the fact that some evolutionary products are not adaptations, the general statement does not substantiate the conditional premise. In sum, then, the conditional premise is false because the evolution of a trait does not entail it being an adaptation—and so far nobody has claimed that consciousness is exempt from this general statement.

Most, if not all, of the examples above thus appear to be unsound deductive arguments that reason from evolution to adaptation with certainty, yet the argument from evolution could work as an inductive argument: Because evolution by itself is not a sufficient condition of adaptations, additional support is needed to strengthen the conclusion. This was already recognised, for example, by William James (1879) who stated that consciousness had an evolutionary origin and who then, instead of simply inferring that consciousness was adaptive, went on to provide evidence for its usefulness. To conclude from its evolution that consciousness is an adaptation, questions such as the following should be answered and supported (see also section 1.2): Is the considered concept or aspect of consciousness at a level of aggregation at which potential adaptations may be found? Was consciousness behaviourally efficacious so that natural selection could act on it? Were conscious organisms fitter than their less conscious competitors? If so, was consciousness actually selected for? And was consciousness heritable so that it could be passed on reliably? The argument from evolution by itself does not establish that consciousness is an adaptation, but additional support, such as the arguments considered in the remainder of this chapter, might warrant this conclusion.

Based on the preceding analysis of the argument from evolution, the argument's promise can be evaluated as described in section 2.4. The argument from evolution aims to show that consciousness is an evolutionary adaptation. I give this argument the higher of the two possible relevance scores for means of inquiry in this category, for three reasons: The argument is applicable to any aspect of consciousness and to any species whose consciousness can be supported as having evolved, it aims to provide definite knowledge, and it specifies the feature of consciousness that allegedly makes it an adaptation (i.e., having evolved). Although the argument highlights this requirement for adaptations, which is in fact a necessary condition for all evolutionary explanations of consciousness, it does not deserve a fertility bonus because it does not carry additional benefits for other approaches (i.e., being an adaptation already entails having evolved). Hence, the relevance of the argument from evolution to the research goal is still only low (3.5).

In evaluating the argument's efficacy, we may take the premises of argument (A) as input to the argument and its inference as procedure. The inference of this deductive argument is valid (10), but the conditional premise is false (0, below the bar), rendering the argument unsound and unpromising. The much less common inductive version of the argument without additional premises fares only slightly better on the input criterion, though it avoids the cutoff for potentially promising means: Having evolved supports the first premise's consequent that consciousness is very likely an adaptation weakly (2). Even though the deductive argument's promise has already been determined, it is worthwhile to evaluate the argument's practicality to enable comparisons and potential improvements. Much good evidence about the second premise of argument (A) is already available or could easily be so within a decade (9), as indicated by the above suggestions of support for the inductive version of argument (P2). Giving the argument requires gathering such support, which implies a relatively low consumption of resources (8), dependent, of course, on the extensiveness and quality of support gathered. The analysis of the argument from evolution examined the underlying reasoning and showed clearly why the argument is unpromising and how it could be improved. Due to space limitations, the following nine evaluations of arguments in this chapter are reported in less detail.

3.2 Arguments based on experience or others

Do you feel that consciousness might have come about as a functionless by-product or by chance alone? It has been claimed that "the standard view is that consciousness evolved because it conferred its bearers an adaptive advantage" (Polger & Flanagan, 2002, p. 21) and similarly that "the dominant view is that consciousness is itself an adaptive feature, playing a crucial functional role" (Hameroff, 1999, p. 245). This section identifies the main reason why nearly everyone who has thought about consciousness is apparently convinced that it is an evolutionary adaptation. Most of the authors who realised that they needed to provide support for such a claim about consciousness, aside from stating that consciousness exists and has evolved (see section 3.1), seem to have relied on introspective insight into their own conscious experience for evidence. After analysing and evaluating the corresponding influential *argument from subjective centrality*, I assess the two quotations given above, which appeal to the prevailing view as an authority in support of consciousness being an adaptation.

Argument from subjective centrality. Introspection and communication with other people show us how central consciousness is to our lives. In fact, it is hard to imagine what our lives would be like without consciousness; it seems to affect our experiences and many of our actions vitally. Because we experience consciousness as such an essential feature of our lives, we tend to infer that natural selection would not have ignored but rather favoured it. For example, Gray (1971) argued that, "although certain minor characteristics of organisms might arise by association with other characteristics themselves possessing survival value, i.e., as epiphenomena, it seems implausible that such a major characteristic as the possession of consciousness could do so" (pp. 252–253). In summarising this inclination to believe that consciousness is an adaptation because of its centrality to us, Ruse (2000) abstracted from introspective evidence to human nature:

> Consciousness seems a very important aspect of human nature. Whatever it may be, consciousness is so much a part of what it is to be human that Darwinians are loath to say that natural selection had no or little role in its production and maintenance. (p. 197)

The suggestion is thus that the subjective importance of consciousness points to consciousness being an adaptation.

The argument from the centrality of consciousness is often presented with more emphasis on the adaptive significance of consciousness than on the action of natural selection, as exemplified by the following three quotations. Århem and Liljenström (1997) called consciousness "an outstanding feature of man, and presumably of other species as well," and explained that, "according to the theory of evolution, outstanding features of organisms have evolved because they have a survival value for the organism" (p. 610). Franklin (2005) also believed that subjective consciousness was "too central a trait ... to have evolved and survived without serving a vital purpose" (p. 118; his other reasons are discussed below). Finally, Baars (1988) asserted that, "like many other major biological phenomena, consciousness plays more than one significant adaptive role" (p. 377). The frequently given centrality argument sounds plausible because we tend to agree with its authors that the lives we lead hinge on us being conscious and, consequently, that being conscious probably helped our ancestors in their survival.

The basic idea that the significance of something for an organism can support the claim that it is an evolutionary adaptation is not a bad one, provided that the something is likely to be an evolved trait of that organism. Plausible examples of evolved traits that are central to our lives are bipedalism, tool use, and parental love. All of the examples of centrality arguments quoted above do assume that consciousness is a product of evolution at the appropriate level of aggregation, so the (inductive) argument from subjective centrality may be phrased as follows:

Consciousness is central to us.
Consciousness has very likely evolved.

Therefore, consciousness is very likely an adaptation.

The evolutionary premise needs to be made explicit because stating the importance of something, by itself, is at most weak support for an adaptation claim. This is because the something could be central, for example, only at a particular time to a single individual. More generally, the central something may be very rare in the population, it may not be heritable, or it could simply be too recent for evolutionary processes to have acted on it. The expanded argument from subjective centrality clears the evolutionary hurdle, but there are other difficulties with it.

One weakness of the argument is that the notion of centrality is not clear. When can something be considered central in the sense that qualifies it as a potential adaptation? Consider a well-known counterexample which shows that seeming significance is not necessarily a good guide for identifying adaptations: The redness of blood looks like an important feature. The colour stands out, in particular when compared to the transparency of other bodily fluids such as saliva, sweat, and tears. Blood could be red, for example, because the redness draws attention to an animal's wound requiring action. Instead, blood is red because it comprises iron-containing haemoglobin used for oxygen transport. Hence the redness of blood is not central. The next question then is how much we need to know about something before we can reliably identify it as central in the relevant sense. Knowing why something is central would make the centrality argument stronger, but the aim of the arguments in this chapter is to offer support for consciousness as an adaptation that does not require knowledge of hypothesised function/s. The vagueness of the intended centrality weakens the link between the centrality premise and the adaptation conclusion.

The basis for the centrality argument about consciousness is, as already mentioned, introspective evidence—principally from one's own introspection, but also in the form of reports about others' introspections. That is to say, we typically know how important consciousness is to our lives from introspection, which itself depends on consciousness: Insofar as *introspection* refers to the conscious action of looking into the consciously accessible parts of one's mind, we cannot introspect unconsciously. It seems quite possible therefore that the conscious nature of introspection could lead us to misjudge the significance of consciousness. This idea is in line with Flanagan's (1992) claim that consciousness makes us overestimate the part it plays in our mental lives. Researchers who have contended that "the individual values his consciousness above all else" (Barlow, 1980, p. 81), that it gives rise to "a Self whose life is worth pursuing" (Humphrey, 2006, p. 131), or similarly that phenomenal consciousness is "the phenomenon that makes life worth living" (Bringsjord et al., 2002, p. 121) certainly value consciousness deeply. Because introspection is probably not an impartial consciousness meter, it is imprudent to rely on introspection unquestioningly. Other reasons for mistrusting introspection are dealt with in section 4.1, and we should take these warnings about introspection seriously when making or assessing claims about consciousness.

Although introspection presumably leads us to overvalue consciousness, our high esteem for consciousness could nevertheless point to an adaptive purpose in evolution. This is not because we are necessarily in accord with natural selection on what is important (we take most adaptations for granted, such as circadian rhythms, bipedalism, and eyes, and may even dislike some potential adaptations, such as pregnancy sickness, sweating, and jealousy), but because there might be a good evolutionary reason why we perceive consciousness as central. What matters to natural selection is whether the felt importance of consciousness influences our actions in an evolutionarily relevant way, as parental love does for instance. A possible example for such an evolutionary function of consciousness is that we may all have "at some point decided to go on living in large part in order to *continue* to be conscious" (Bringsjord et al., 2002, p. 111). If the subjective importance of consciousness is an inherent feature of evolved consciousness and if it is central to our lives, then valuing consciousness itself might call for an adaptation explanation by the above centrality logic (e.g., Barlow, 1980). It is thus possible that we value consciousness because valuing consciousness contributes to our survival, but it is unlikely that we value consciousness because it is an adaptation.

To recapitulate the main points about the argument from subjective centrality, virtually everyone today feels that consciousness is important: “For us, consciousness is, to put it barbarically, a big deal” (Bringsjord et al., 2002, p. 111). Such claims are based on introspection in the first instance: Consciousness *feels* central to us. However, the subjective significance of consciousness assists adaptation claims about consciousness only insufficiently, because consciousness could be less central than introspection makes us believe and because our own appreciation of our features does not indicate their adaptation status reliably. Generalised statements, such as that consciousness is an adaptation because it is a central characteristic of how we live or of human nature, appear to rely less on introspection but also fail to forge a strong connection between centrality and adaptation status. Whether or not consciousness is said to be central according to our values, which may be irrelevant to natural selection, it must enable, or at least help to enable, evolutionarily important behaviours. If generalised arguments from the centrality of consciousness were claiming the latter, they would seem to involve unspecified assumptions about the function/s of consciousness; this would make an adequate evaluation difficult, and not belong to the present chapter.

The evaluation tool serves to determine the promise of the argument from subjective centrality. Like the argument from evolution in the previous section, this argument receives the higher possible relevance score (3.5) for means of inquiry that aim to support consciousness as an adaptation: For now the argument may only be usable for human consciousness, but it addresses a specific characteristic of consciousness and gives a clue why consciousness might be an adaptation (i.e., its subjective centrality). The clue is, however, too vague for other approaches to benefit (e.g., when and why is centrality evolutionarily relevant), so the argument does not qualify for a fertility bonus. The evidence for the premises is high (8), though it would be worth providing support for the subjective centrality of consciousness in other current and in earlier cultures. Because of the problems with overvaluing consciousness and simplistically equating our concerns with those of natural selection, the argument’s premises support the conclusion only slightly more (3) than the premises of the inductive argument from evolution. The availability of good data for applying the argument to human consciousness is very high (9), and the resource cost low (7; cf. the argument from evolution, with an additional premise to be supported), depending again on what support is collected. Taken together, the argument is moderately promising (5.2).

Argument from the standard view. Arguments from authority for consciousness being an evolutionary adaptation are rare in the consciousness literature. This is not surprising as there are only few people, if any at all, who would count as genuine experts on this topic. Claims that it is the standard or dominant view that consciousness is an adaptation were already quoted at the beginning of the present section. It is not clear who the authority is on which these arguments rely. If it is the part of the general population who has a basic knowledge of the theory of evolution, I agree with the assessment of the mainstream view on this matter, yet question its significance. It is doubtful that the general population is informed enough to act as an authority on the adaptation status of consciousness. Their standard reasoning might go like this: Consciousness is useful because it is central to me, so it probably is an adaptation. Both inference steps could be challenged, but for now it suffices to refer back to our likely bias regarding the centrality of consciousness and to the limited inferences we can draw from any trait's subjective centrality to its adaptive usefulness. Most of us are neither qualified nor impartial enough to be appealed to convincingly as an authority on whether consciousness is an adaptation.

Instead of invoking folk opinion on the adaptation status of consciousness, Hameroff (1999) and Flanagan and Polger (Flanagan, 1999/2000; Polger & Flanagan, 2002) might rather have had the mainstream view of consciousness researchers in mind:

> The standard view among qualified experts is the best available approximation to the truth.
> The standard view among consciousness researchers is that consciousness is an adaptation (or adaptive).
> ___
> Therefore, consciousness is probably an adaptation (or adaptive).

This interpretation of the argument is supported by Hameroff's quotation being part of his introductory remarks to a section in the proceedings of a consciousness conference (and as such it is probably primarily meant as a description, not an argument). The following quotations also suggest the consciousness-researcher interpretation for Flanagan (1999/2000): "It is not completely uncontroversial, but it is pretty close to being uncontroversial that sensory consciousness is an adaptation" (p. 102); "Darwin called the eyes 'organs of extreme perfection.' There is little doubt that visual consciousness is an adaptation" (p. 109). Controversies about the adaptation status of sensory or visual consciousness would typically occur among researchers, and the quotation of Darwin reinforces the focus on researchers. The question then is whether there is any merit in engaging consciousness researchers as authorities on the adaptation status of consciousness.

The majority of consciousness researchers believes that conscious vision and sensation, and consciousness more generally, are evolutionary adaptations; here I concur with Hameroff (1999) and Flanagan and Polger (Flanagan, 1999/2000; Polger & Flanagan, 2002). However, even if the researchers' knowledge was sufficiently impartial and based on established results of scientific research, their status of authority could still be undermined by a lack of adequate scientific evidence. Chapters 3 to 5 give an indication of the currently available evidence for consciousness being an adaptation. This evidence is limited and, in addition, unlikely to be known comprehensively by the majority of consciousness researchers. Note also that the apparent usefulness or adaptiveness of certain kinds of consciousness, such as conscious vision and sensation, has been challenged on the grounds that they do not seem to be necessary for the corresponding mental processing and behaviours to occur (e.g., Jaynes, 1976/2000; Velmans, 1991; see sections 1.3 and 1.4). Nevertheless, the present argument from expert authority looks more promising than its folk version, especially once more scientific evidence is available and taken into account.

To be evaluated with the tool is thus the promise of the inductive argument from the standard view among consciousness researchers. Because the argument does not provide any additional information about consciousness other than that it might be an adaptation, its relevance is lower (3) than that of the two arguments evaluated so far. The first premise expresses an acceptable general argument from authority, and the second premise is, overall, also well-supported (7): Although many consciousness researchers have not publicly addressed the issue, the most frequently expressed view is that consciousness is an adaptation (a small number of researchers who hold otherwise are cited in the next section). Given that consciousness researchers may be only partially knowledgeable about evolutionary theory and susceptible to bias about consciousness, the support for the conclusion might be moderate (4); it would be higher if the authority consisted of experts on the evolution of consciousness. Evidence regarding the second premise is already available and could easily increase (e.g., dedicated conference symposia, special journal issues; 8). Surveying the relevant academic literature for standard-view support requires comparatively little time and hardly any money (9). On the whole, the argument is about as promising (5.1) as the argument from subjective centrality. Remarkably, all subevaluation scores of these two inductive arguments are within 1 point of each other, and they both appeal to nonevolutionary authorities to support the claim that consciousness is an adaptation.

3.3 Five simple adaptation arguments

At least on the surface, evolutionary reasons would appear to be more suitable for making the case that consciousness is an evolutionary adaptation than introspection-based or standard-view arguments. I discuss five such reasons in this section, which are at least loosely connected with evolutionary theory: Among the most direct indicators of the adaptation status of consciousness may be such features as its apparent adaptedness and biological cost. Another argument infers that phenomenal consciousness must be an adaptation from an evolutionary history constructed to explain its nature. Then there is the observation that support for consciousness not being an adaptation is lacking. And finally, people sometimes generalise from the adaptation status of particular kinds of consciousness to that of consciousness in general. Most of these arguments for consciousness being an adaptation, which I examine and evaluate in turn, have the character of passing comments—maybe because no opposition was expected.

Argument from apparent adaptedness. If consciousness has been selected for, conscious organisms must have been better adapted to the selective environment than their less conscious counterparts. Consequently, some researchers have argued that if consciousness appears to render its possessors better adapted, it is more likely than not an adaptation (assuming it has evolved): "Given its apparent adaptedness, the null hypothesis must be that consciousness was selected for in the process of evolution by natural selection" (Polger & Flanagan, 2002, p. 30).[2] Using the more common term *adaptive* for traits, the argument can be restated as follows:

[2]Incidentally, if we draw on the usefulness of consciousness to support the claim that consciousness is an adaptation, we should not use the adaptation statement to infer that consciousness is functional, as its efficacy has already been assumed at the outset.

An adaptive trait is probably an adaptation.
Consciousness appears adaptive.

Therefore, consciousness is probably an adaptation.

Taking a trait's adaptiveness as tentative support for it being an adaptation seems reasonable in the absence of agreed-upon estimates of the prevalence of adaptations compared to other evolutionary products: Although there are other causal processes which could have led to the trait being adaptive (and the corresponding alternative explanations should be examined as well), the simplest or most economical explanation of its adaptiveness is that it is an adaptation.

What makes the application of the evaluation tool to this argument difficult is that it is not clear what the basis is for the claim that consciousness is apparently adaptive, nor even what its basis should be. On the one hand, the authors cited above asserted that (consciousness of) acute pain is, prima facie, useful for minimising injury. On the other hand, they acknowledged that "although it seems likely that some varieties of consciousness are adaptations, specifying what the adaptive advantage of a kind of consciousness *might be* is difficult" (Polger & Flanagan, 2002, p. 27), thereby suggesting that the adaptation judgement need not depend on knowledge of evolutionary function/s. For this reason, I do attempt to evaluate the argument from apparent adaptedness here.

Starting again with the argument's relevance to the research goal, I assign the lower possible relevance score to the present argument (3) despite its similarity with the argument from evolution in this respect. Its potential contribution is lower than that of the deductive argument from evolution because it does not aim to provide definite knowledge; rather, it merely seems to recommend the research strategy, methodological adaptationism, pursued by all arguments in this chapter. The evidence concerning the apparent adaptiveness of consciousness—possibly from intuition, introspection, or personal experience—is not strong, but in its favour (4). However, apparent adaptiveness is only weak adaptation support (3), as it may well differ from actual adaptiveness: Many intuitive beliefs about consciousness have already been overturned, for instance, by evidence for unconscious mental processes. The availability of the aforementioned data is probably high (8), and the resource cost incurred by collecting this data—whether from participants or the existing literature—low (8). The argument from apparent adaptedness by itself is moderately promising (4.2).

Argument from biological cost. Another feature of consciousness, its biological cost, may suggest that it is an adaptation. Searle (2001) relied on such reasoning: "In humans and higher animals an enormous biological price is paid for conscious decision making ... To suppose that this plays no role in inclusive fitness is ... like supposing that vision or digestion played no evolutionary role" (p. 509); as did Baars (1993): "By any reasonable measure, the conscious stream is biologically very expensive. ... What evolutionary benefits could justify such costs?" (pp. 285–286; see also Keenan, Rubio, Racioppi, Johnson, & Barnacz, 2005). Once again it is helpful to spell out the premises and the conclusion of the argument:

A biologically costly trait is probably an adaptation.
Consciousness is biologically costly.

Therefore, consciousness is probably an adaptation.

The biological costs of (particular kinds of) consciousness referred to in this argument include the energy consumption of its neural bases (Baars, 1993; Keenan et al., 2005; Searle, 2001), the consequences of being limited to a single conscious stream (Baars, 1993), and the investments in rearing conscious offspring (Searle, 2001).

One way in which the argument from biological cost can go wrong is illustrated by the following example. Beginning an argument by contradiction, Franklin (2005) assumed that subjective consciousness was a by-product. He had previously granted that a particular function of subjective consciousness could in principle be performed unconsciously. Subjective consciousness would then have to be generated by a mechanism that was separate from the mechanism responsible for performing its function, and hence, would carry its own computational or biological costs. He concluded that because subjective consciousness had evolved in spite of these costs, it had an adaptive significance outweighing them. A problem with this argument is that a by-product is a by-product of an adaptation and is therefore not produced by a separate mechanism—think of the redness of blood or the belly button as examples. Because a costly by-product cannot be selected against while its adaptation is being selected for, their costs and benefits need to be considered jointly. In general, costliness can serve to identify adaptations (Andrews et al., 2002), in that a costly trait that could be selected against would not normally evolve unless it was associated with benefits, either directly (adaptation) or indirectly (by-product of an adaptation).

The argument from biological cost performs well on the relevance dimension of the evaluation tool (3.5) compared to other means of inquiry in the same category, because it specifies why consciousness is probably an adaptation; yet it does not entail additional benefits for other approaches to the research goal. The conclusion of the argument follows from the premises (10), whose support seems closer to *high* than *moderate* (7): The first premise is reasonable if the trait has evolved, though the proportion of costly adaptations to costly by-products is not well-established. Support for the second premise comes from, for example, evidence for lower neural activity both during lower levels of consciousness (e.g., Laureys, Owen, & Schiff, 2004; Shulman, Hyder, & Rothman, 2009) and for less conscious contents (e.g., Dehaene et al., 2001; Tong, Nakayama, Vaughan, & Kanwisher, 1998). Although we should not expect all confounds to be removed from measurements of the cost of consciousness within the next decade, the accessibility of data about the premises is high (7). Especially if a large proportion of the data is or becomes available as part of other studies, the argument's resource cost is low (7), and thus its promise for contributing to an evolutionary explanation of consciousness is moderate to high (6.2).

Argument from evolutionary history. If we knew in what stages consciousness evolved, we might be able to infer whether natural selection was involved in any of them. In reply to self-posed questions about the nature of phenomenal consciousness, Humphrey (2008) summarised his theory of conscious sensation, according to which sensory responses became internalised during evolution. Before addressing the evolutionary function of phenomenal consciousness, he stated that "we can surely assume that the kind of development I have sketched above will not have happened accidentally. It must be the result of natural selection favouring genes that underwrite the specialised neural circuits" (p. 267). A more formal version of the argument is the following:

If the evolutionary history of changes in a trait necessitates that the trait was selected for, the trait is, or was, an adaptation.
The evolutionary history of consciousness necessitates that it was selected for.

Therefore, consciousness is, or was, an adaptation.

Humphrey's strategy of inferring that consciousness is an adaptation from its evolution is appealing, but difficult to implement successfully: One would need a well-supported account of the evolutionary development of consciousness without any reference to the action of natural selection—no likely selection pressures, function/s of consciousness, or evidence about its adaptation status.

Although the present evaluation is of means of inquiry, not instances of their use, it might be instructive to see how well Humphrey's (2008) argument worked. Unfortunately, his theory of the evolution of phenomenal consciousness appears to assume natural selection: "Sentition [the activity of sensing] has been subtly shaped in the course of evolution so as to make our picture of it have those added dimensions of phenomenality" (p. 266). Other publications of his confirm this reading. For example, Humphrey (2000) declared consciousness being an adaptation his default assumption and set out to identify the evolutionary function of conscious sensory experience. He combined the ideas that the subjective quality of sensations was selected for and that it is private, hence invisible to evolution, by suggesting that it became so only after its selection. Humphrey then outlined his theory of the evolution of sensation and concluded that "we can both make good on our ambition, as Darwinists, to explain sensory quality as a product of selection, and we can accept the common sense idea that sensations are as private as they seem to be" (p. 250). In fact, natural selection was already part of Humphrey's (1992) evolutionary history of sensation when he first began to develop his theory of sensation. Overall then, Humphrey (2008) did not implement his own strategy, and his actual approach (1992, 2000) does not belong to the present chapter because his evolutionary history involves hypothesised functions of consciousness.

For means of inquiry aiming to show that consciousness is or was an adaptation, the proposed argument from evolutionary history is relatively relevant to the research goal (3.5) because it can be expected to identify changes involved in the emergence of the adaptation. By specifying the evolutionary history of consciousness, the argument carries significant additional benefits (+1), especially for the origin-and-history approach. I find no fault with the inference (10), and the first premise is only phrased too strongly. However, so far there is at most weak evidence for the second premise (5); almost all hypothesised evolutionary histories of consciousness do involve changes due to natural selection, but this could, for example, reflect a shared bias instead of the truth. The Achilles heel of the argument, at least in this evaluation round, is the very low availability of data for the second premise within a decade (1, below the bar); we still seem to be far from a well-supported nonselective evolutionary history of consciousness.[3] The cost of the resources required to reconstruct as much as possible of the evolution of (maybe just one kind of) consciousness is high (2), which is why it appears unwise to neglect potentially valuable selective evidence in the face of a largely missing historical record. In sum, the argument is currently unpromising due to insufficient data but worth including in future evaluation rounds.

Argument from the absence of opposing evidence. Another argument centres on the idea that consciousness is more likely to be an adaptation not only if there is support for it being an adaptation but also if, in addition, there is little support for it not being one:

> If the evidence about a trait as an adaptation is mostly supportive, the trait is probably an adaptation.
> The evidence about consciousness as an adaptation is mostly supportive.
> ---
> Therefore, consciousness is probably an adaptation.

For evidence to be mostly supportive, there needs to be clearly more evidence for the trait being an adaptation than opposing evidence, such as evidence against the trait being an adaptation or evidence for the trait being some kind of nonadaptation. The argument has been used by Franklin (2005) who, besides offering positive reasons for thinking that subjective consciousness is an adaptation, also gave the negative reason that “either an argument, or evidence, for subjective consciousness being a byproduct would be needed [to support the anti-adaptation case]. I know of neither” (p. 118). The negative reason is problematic in this instance because Franklin stated his ignorance in reply to a self-posed question that contained a suggestion for such an argument: “It is hard to see why it [consciousness] is not just a byproduct, given that we can explain the behavior purely in terms of the computational mechanisms” (p. 117). Regardless, the general argument depends on how much opposing evidence there is.

Hardly anyone has seriously claimed that consciousness is an evolutionary product other than an adaptation. One exception is Thomas Huxley (1874/1882) who argued

[3] Leary and Buttermore (2003) have attempted a detailed reconstruction of the history of self-reflection since the last common ancestor of humans and chimpanzees, but they cautioned that “far more research is needed to offer a cogent and convincing explanation of *why* [emphasis added] symbolic ability and the conceptual self evolved precisely as they did” (p. 395).

that consciousness was “a collateral product” (p. 236) of brain processes without any causal influence on such processes. He reasoned that our consciousness was preceded in evolution by simpler forms of consciousness in nonhuman animals, but failed to consider how such an epiphenomenal consciousness might have evolved. This oversight could be remedied by Blakemore’s (2008) suggestion that consciousness evolved as a by-product of the ability for metarepresentation, if we could explain the latter’s association with consciousness. Note that the coupling of Blakemore’s metarepresentations with conceptual language entails that most animal species are not conscious. Likewise, Macphail (1998, 2000) speculated in his comparatively well-developed theory of the evolution of consciousness that only the evolution of conceptual language brought about consciousness in humans. He construed the benefits of self-consciousness, namely helping to direct thoughts and to predict others’ actions, as possible additional advantages of language. Because the experience of pleasure and pain is not necessary for appropriate behaviour to occur, Macphail could not conceive a function for phenomenal consciousness and concluded that it was an evolutionary by-product of self-consciousness. In doing so he erroneously inferred etiological epiphenomenalism from inessentialism about phenomenal consciousness (see section 1.4) and overgeneralised from pleasure and pain to other forms of phenomenal consciousness.

Two other nonadaptation accounts of consciousness deserve mention. The thrust of Gould’s (1991) argument, as I understand it, is that “complex human consciousness” (p. 59) evolved and is presently useful but did not originate as an adaptation, just like almost everything else human brains, which did become enlarged for adaptive reasons, do today. This argument is rooted in his belief that, “surely, for something so complex and so replete with latent capacity as the human brain, spandrels must vastly outnumber original reasons, and exaptations of the brain must greatly exceed adaptations by orders of magnitude” (p. 57).[4] Most relevant here is that Gould did not provide additional support for his claim and that he was only concerned with the nonadaptive origin of human consciousness, not with its subsequent evolutionary development nor with other kinds of consciousness. According to Rosenthal (2008), consciousness of thoughts and desires occurs once self-ascriptions of thoughts and desires as causes of behaviour (based on other- and self-observation) have become automatised through

[4] In this quotation the term *spandrels* most likely refers to co-opted by-products and *exaptations* to co-opted adaptations. Gould’s (1991) use of the terms was inconsistent, including the question of whether these traits’ current utility has biological significance (Buss et al., 1998).

practice, but it does not add any significant benefit. This explanation is based on "the difficulty of finding any credible function for the consciousness of psychological states" (p. 837) and assumes Rosenthal's higher-order-thought theory. Support for consciousness not being an adaptation could also come from deflationary evolutionary explanations (see section 5.1) and other epiphenomenalist arguments (see section 1.3), but the latter have so far only been made in nonevolutionary contexts and are inconclusive.

Returning to the argument from the absence of opposing evidence to evaluate it, its relevance to the research goal merits only the lower score in the respective category of means of inquiry (3), because available evidence is reviewed merely for its direction. A comparison of the digression in the last two paragraphs with Chapters 3 to 5 illustrates that most evidence is indeed for consciousness being an adaptation. However, the other (generously phrased) premise is less certain (5), mainly because the general principle is applied to potential adaptations and, furthermore, meant to be applicable to consciousness: The near absence of nonadaptation evidence about consciousness can alternatively be explained, at least in part, by the prevalence of methodological adaptationism or the existence of a general bias in favour of consciousness as an adaptation. This bias has been proclaimed by several researchers as the default position (e.g., Humphrey, 2000; Polger & Flanagan, 2002) and is likely to discourage researchers to look for and identify opposing evidence. Apart from this significant concern, the conclusion follows from the premises (10); and determining whether there is more evidence for than against the hypothesis that consciousness is an adaptation is very resource inexpensive (9). Higher-quality data could be available within a decade (6), but would require more efforts dedicated to identifying opposing evidence, which goes against the continuing adaptation tendencies. The synthesis procedure indicates that the argument is moderately promising for increasing our knowledge about the biological evolution of consciousness (5.7).

Generalisation argument. Some adaptation claims are about consciousness in general, whereas others are about particular kinds of consciousness. Consider the phrase preceding an argument from the standard view: "Generalizing from cases like vision and pain, the standard view is that consciousness evolved because it conferred its bearers an adaptive advantage" (Flanagan, 1999/2000, p. 109). The suggestion here is that the adaptation status of visual and pain consciousness extends to consciousness in general:

If some subtypes of a trait are adaptations, the trait is an adaptation.
Some subtypes of consciousness are probably adaptations.

Therefore, consciousness is probably an adaptation.

The argument raises several issues: How do we know which subtypes to consider? Conscious vision and pain may be adaptations, but that is unlikely for other kinds of consciousness such as conscious schizoid thinking (Flanagan, 1999/2000), conscious chronic pain, and dreaming (see section 5.1). Moreover, it does not follow necessarily from parts of a structure or function being adaptations that the whole structure or function is an adaptation (Flanagan, 1999/2000, himself makes this point in a somewhat different context). Assuming that some kinds of consciousness are adaptations, it is unclear what the encompassing consciousness is, how it has evolved in relation to its subtypes, and hence whether, and how, its adaptation status and those of its kinds are related in a general way.

This generalisation argument, the final means of inquiry to be evaluated in this section, receives the lower possible relevance score of means for showing that consciousness is an adaptation (3): The relation of consciousness to its subtypes that supposedly makes it an adaptation may simply follow the alleged general rule in the first premise without pointing to any evolutionarily relevant feature of consciousness. It is this first premise that breaks the argument (0, below the bar), as should be evident from the aforementioned issues, especially the consequent not following logically from the antecedent. Additionally, the first premise very much depends on our ability to correctly identify traits and their subtypes at appropriate levels of aggregation. In my phrasing of the argument, the inference is valid (10). Further, we can expect the availability of provisional evidence about the adaptation status of different kinds of consciousness to be high within a decade (7). Gathering this data may require few resources, but might not be sufficient and would in any case need to be gathered for more than one kind of consciousness (7). However, this is not advisable for the purpose of the present argument as it is unpromising due to its false first premise.

The five arguments in this section, despite having a closer connection with evolutionary theory than those in the previous section, are not on average more promising. Two arguments are even unpromising, the argument from evolutionary history because of insufficient data within a decade and the generalisation argument because of a false premise. The arguments that do perform better in the evaluation of promise, the argument from biological cost and the argument from the absence of opposing evidence, can also easily go awry, as illustrated above. In addition, the basis of some of the arguments is not clear, specifically of the argument from apparent adaptedness and the generalisation argument. I suspect that the elaboration of the arguments, which might have revealed problems with them or with instances of their use, either did not occur to the respective authors or appeared superfluous because of the general bias for consciousness being an evolutionary adaptation.

3.4 Evolutionary arguments taking us further

Proper evidence for the adaptation status of consciousness should ideally come from *proper* evolutionary biology or psychology methods. Variations of two such means of inquiry, namely of the comparative method and the special design standard (sometimes taken to include costliness; e.g., Andrews et al., 2002), have been employed as arguments for consciousness having evolved by natural selection. As long as consciousness does not exist exclusively in humans, we might be able to learn more about its adaptation status based on which other species are conscious. Accordingly, the argument from species comparisons relies on animal consciousness to infer that consciousness is probably an adaptation. The argument from complex design, which is introduced in more detail below, bears on the observation that some biological traits appear so complex and well-designed that they are extremely unlikely to have arisen by chance. I now analyse and evaluate the promise of these two arguments.

Argument from species comparisons. The common denominator of the first set of arguments in this section, concerned with the distribution of consciousness in biological organisms, is what I call the *argument from species comparisons*:

> A trait that is present in several animal species is probably an adaptation.
> Consciousness is present in several animal species.
> ___
> Therefore, consciousness is probably an adaptation.

An early illustration of such reasoning with a historical perspective is William James's (1879) statement that "consciousness … has been slowly evolved in the animal series, and resembles in this all organs that have a use" (p. 3). James (1890/1907) later elaborated on this:

> It is very generally admitted, though the point would be hard to prove, that consciousness grows the more complex and intense the higher we rise in the animal kingdom. That of a man must exceed that of an oyster. From this point of view it seems an organ, superadded to the other organs which maintain the animal in the struggle for existence; and the presumption of course is that it helps him in some way in the struggle, just as they do. (p. 138)

James's (1890/1907) argument has intuitive appeal (perhaps excluding his claim to oyster consciousness), but several points—about the general argument, James's version, and another version—need to be examined more closely.

First of all, to provide comparative arguments about consciousness, we must be able to determine whether other biological organisms are conscious. By this I do not mean to bring up the philosophical problem of other minds, which I shelved in section 1.1, but rather a practical difficulty with animal consciousness: Nonhuman animals cannot communicate with us about whether they are conscious or what they are experiencing, by using a conceptual language. However, evidence for their consciousness can, for instance, come from behavioural and neural similarities between them and us and from nonverbal reports. Examples of similarities used in reasoning by analogy are pain responses, efficacy of anaesthetics, brain structures, sensory systems, susceptibility to visual illusions, and cognitive abilities (Allen, 2010). To give a specific example of the use of nonverbal reports, macaque monkeys in Myerson, Miezin, and Allman's (1981) study of binocular rivalry were trained to indicate the direction in which vertical gratings moved by pressing keys. When the direction of movement was opposite for each eye, the monkeys reported direction alternations with similar characteristics (e.g., distribution of durations, velocity dependence) to those reported by human observers in the same experimental situation. Myerson et al. (1981) took this finding as evidence of similar binocular-rivalry experiences in monkeys and humans.

Similarity support for animal consciousness can be boosted by taking possible common evolutionary descent into account (Allen, 2010). If species have a common ancestor and show similar behaviour, the more parsimonious explanation is that the same, not different, mechanisms are involved in producing the behaviour in question, with parsimony reflecting Fisherian likelihood (i.e., the same-mechanism explanation makes the observed similarities more probable; Sober, 2000). Arguments from homology, that is, from similarities due to shared ancestry, require knowledge of what behaviours (and neural processes or structures) are associated with consciousness, as does plain reasoning by analogy (for sample criteria of animal consciousness, see Mesulam, 1998, and A. K. Seth, Baars, & Edelman, 2005). Any comparative argument should be based on evidence that the animals under consideration are conscious.

Granting that some animals are most likely conscious, are there alternative explanations of the distribution of consciousness described by James (1890/1907)? It might be possible, for example, that consciousness spread in an ancestral population common to all of today's conscious species by genetic drift and was then neither directly selected for nor against, because it did not affect its bearers' fitness. Yet this scenario makes it unlikely that consciousness became more complex over time. In accordance with James's claims about the signs of this complexity, human behaviours are generally more flexible and human brains more complex (e.g., encephalisation; Jerison, 1985; number of neurons and synapses; Roth, 2001), particularly so when compared to the cerebral ganglia of oysters. A different alternative explanation is that consciousness spread as the by-product of an adaptation. In this case, consciousness could be associated with the complexity of the brain, which may have increased during evolution for unrelated reasons (Gould, 1985; these reasons might be adaptive, but the original adaptations are likely to be very few and unidentifiable today; Gould, 1991). However, the by-product scenario becomes much less likely once the complexity of consciousness is properly taken into account in the next section. Yet without this complexity, James's comparative argument would exclude neither alternative explanation—a clear weakness, but not a fatal one, as even the anti-adaptationist Gould (1991) admitted that "those characteristics that we share with other closely related species are most likely to be conventional adaptations" (p. 61).

Species comparisons can be used to argue that consciousness is an adaptation either from homology, as James (1879, 1890/1907) did and A. K. Seth et al. (2005) suggested, or from analogy.[5] The notion behind arguments from analogy is that consciousness may have evolved independently in different lineages because of similar selection pressures (i.e., convergent evolution). Or, with Flanagan (1992) quoting Calvin (1991), evolution tends to "reinvent" biologically useful traits, as it has done with powered flight at least four times and with photoreceptors over 40 times (current estimates

[5] *Analogy* refers here to the similarity of functions performed by separately evolved structures and has nothing to do with the aforementioned reasoning by analogy for nonhuman consciousness.

of the latter are much lower; Fernald, 2008; Nilsson, 2005). Flanagan recommended the argument from analogy as a "credible way to argue" (p. 133) for the adaptation status of subjective consciousness, provided that some distantly related species were supported as being conscious; his examples were snakes, salamanders, fish, birds, bats, and dolphins. To be successful, the argument from analogy needs support both for the presence of consciousness in the animal species in question and for its absence in their common ancestor, or at least in today's intermediate species.

Although the comparative method is considered to be useful for identifying mental adaptations in general (Andrews et al., 2002; Grantham & Nichols, 1999), Nichols and Grantham (2000) claimed it was "poorly suited" and "currently of no use for determining whether phenomenal consciousness is an adaptation" (p. 653). They argued that the allegedly common assumption of philosophers that all vertebrates are phenomenally conscious precludes the necessary identification of differences in the consciousness of closely related species. Their only example, Flanagan's (1992) argument, lists only vertebrate species, yet these are listed as distantly related. In other words, Nichols and Grantham criticised arguments from homology, whereas Flanagan argued from analogy. Nonetheless, their criticism is valuable in highlighting the importance of specifying, in arguments from analogy, which intermediate species are probably unconscious (or differently conscious). In addition, their criticism suggests studying the potential adaptive divergence of consciousness, that is, exploring how differences in consciousness between closely related species correlate with their ecological niches.

The evaluation of the argument from species comparisons applies to its different versions. Although the argument cannot be used on distinctly human kinds of consciousness, it is more relevant to the research goal than some other means of inquiry in the an-adaptation category, as it identifies animal species in which consciousness is probably an adaptation (3.5). Depending on how much of the relevant species similarities is revealed, the argument is likely to provide additional leads for other approaches (+0.5). The first premise does not completely exclude alternative explanations, and would not do so even if it stated the particular version of the argument being used; this premise would be stronger if not only the presence but also the absence (or difference) of consciousness in closely or distantly related species were specified (7). As in previous arguments, I have expressed the premises so that they support the conclusion (10). Data about animal consciousness can be difficult to collect but is, or could soon be, available (see D. B. Edelman & Seth, 2009; A. K. Seth et al., 2005), albeit to differing degrees for the different kinds of consciousness, their operationalisations, and animal species (8). These factors also influence the argument's resource cost, as does the need to include evidence about several animal species (7). Overall, the promise of the argument from species comparisons qualifies just as *high* (6.5).

Argument from complex design. A crucial element of what makes evolutionary adaptations special is their *special design*: Natural selection has shaped each adaptation to better solve a problem posed by the organism's environment. A trait's features of special design can therefore indicate that it is an adaptation (Tooby & Cosmides, 1992; Williams, 1966); special design has even been considered the "leading evidentiary standard" (Andrews et al., 2002, p. 496) for identifying adaptations. Among special design features, it is complexity that "provides some of the most compelling evidence" (Polger, 2007, p. 72) for adaptations. Not all adaptations are complex, but if a biological trait is complex, "we can reasonably expect that the trait was formed by natural selection" (Polger, 2007, p. 80). This is because natural selection is the most likely—many have claimed the only (e.g., Dawkins, 1985; Futuyma, 2005; Gould, 1997; Pinker, 1997; Stearns & Hoekstra, 2005)—natural process that can produce biological traits with complex functional design. The corresponding argument about consciousness goes as follows:

> A trait that exhibits complex design is probably an adaptation.
> Consciousness exhibits complex design.
> ___
> Therefore, consciousness is probably an adaptation.

The remainder of this section deals with an argument from evolutionary design and an argument from structural complexity (roughly the respective authors' terms), both of which have been advanced to support the view that phenomenal consciousness is an adaptation.

In presenting his argument from design, Flanagan (1992) contended that subjective consciousness "has been hooked onto many distinct mental systems" (p. 134) nonrandomly. He suggested two ways in which consciousness could have evolved, depending on the cognitive model employed: Either consciousness is domain-specific and had to be reinvented for each distinct sensory module and for the declarative memory system, or there is a single consciousness that had to be connected to each module. Either way, consciousness may be considered complex because cognitive models suggest structures that connect conscious experience with the senses and memory. These connections in turn explain why neuropsychological deficits can affect specific parts of consciousness. Flanagan pointed to isolated deficits in subjective consciousness, such as blindness, colour blindness, deafness, and prosopagnosia (i.e., the inability to recognise faces). He also claimed that the design of subjective consciousness is nonrandom because we have conscious access to "primarily just what we have the most need to know about: conditions in the sensory environment, and past facts, and events" (p. 134). The apparent relevance of conscious information to our successful survival and reproduction is supposed to help create a picture of subjective consciousness as complex, specialised, and evolutionarily relevant.

Even though we are not concerned with epiphenomenalist worries here, I want to raise three points about a criticism of design arguments such as Flanagan's (1992). Nichols and Grantham (2000) dismissed what they call "functional complexity arguments for phenomenal consciousness" (p. 656) as begging the question against epiphenomenalism: The alleged fallacy is using an assumed function of consciousness to show appropriate functional design. My first point is that Flanagan did not state what the biological advantage of conscious over unconscious information might be, nor does his complexity defence depend on his last claim that consciousness appears adaptive. Second, matching features of biological traits with expected design features is a well-known evolutionary method called *reverse engineering* (e.g., Lewontin, 1978; Tooby & Cosmides, 1992; Williams, 1966), which is admittedly not suited for fighting epiphenomenalism. Third, Nichols and Grantham's own complexity argument also failed to defeat epiphenomenalism, suggesting that such arguments are generally unsuitable for this purpose (Polger, 2007). Complexity arguments might, however, be suitable for supporting consciousness as an adaptation, which I determine in their evaluation following the description of Nichols and Grantham's complexity argument.

When looking to construct an argument from the complexity of consciousness, it makes sense to study examples of complexity arguments in biology. Sometimes biologists are convinced that a biological trait is an adaptation, even if its usefulness is initially unclear, because the trait shows structural complexity (Williams, 1966). This was the case for the lateral line system in fish, which subsequent evidence strongly indicated is an adaptation. Nichols and Grantham (2000) maintained that phenomenal consciousness is structurally complex in the same way as the lateral lines are. The lateral lines are made up of mechanoreceptors along a fish's length that connect the lateral line nerve with the lateral line canal just underneath the skin. The analogous structure consists of "several independent input channels that feed into a more central mechanism" (p. 663). Nichols and Grantham cited the philosopher Searle (1992) in support of the claim that consciousness involves the unification of experiences across short time spans as well as within and across different modalities, and pointed to selective neuropsychological deficits of consciousness to demonstrate the existence of multiple independent input channels to consciousness. They added that the presence of consciousness in humans shows that it has neither been selected against nor deteriorated, as the lateral lines in some fish species have in which they had become useless. Hence, consciousness appears structurally complex and adaptive.

Compared to other instances of the use of arguments in the present chapter, Nichols and Grantham's (2000) argument stands out as the most developed and is worth pointing out as a model. The length of their argument already indicates this: a core of nine journal pages versus a single sentence in many cases (Flanagan's, 1992, comparative and design arguments span a paragraph and a book page respectively). Nichols and Grantham explained their hypothesis in detail, gave empirical evidence for it, addressed disanalogies between the lateral lines and phenomenal consciousness, considered possible objections, and even asked biologists whether they expected the phenomenal consciousness system, abstractly described, to be an adaptation. The apt use of a biological analogy strengthens their argument, but the high level of abstraction of the analogous structure (i.e., a central mechanism with independent input channels) detracts from the benefit gained. Regardless, the argument from complexity presumably looked sufficiently promising for Nichols and Grantham to develop it in some detail for a particular case.

As before, I apply the evaluation tool to the general argument—here from complex design—not to specific instances of its use (e.g., Flanagan, 1992; Nichols & Grantham, 2000; Polger, 2007). The argument from complex design clearly deserves the higher possible relevance score for means of inquiry with the same aim (3.5): It identifies a feature of consciousness that characterises it as an evolutionary adaptation, and this may additionally provide clues about the function/s of consciousness (Nichols & Grantham, 2000) and its evolution (+0.5). The argument's premises fully support the conclusion (10), but the weaker second premise has currently only moderate support (6)—even Nichols and Grantham (2000) speak of *apparent complexity.* Nevertheless, there is much data about several potentially complex kinds of consciousness

Table 3.1
Subevaluations of Means of Inquiry That Aim to Show That Consciousness Is (or Was) an Evolutionary Adaptation

			Efficacy		Practicality		
§	Means of inquiry	Relevance	Input	Procedure	Avail.	Cost	Promise
3.1	Evolution	3.5	~~0~~	10	9	8	~~5.1~~
3.2	Subjective centrality	3.5	8	3	9	7	5.2
3.2	Standard view	3	7	4	8	9	5.1
3.3	Apparent adaptedness	3	4	3	8	8	4.2
3.3	Biological cost	3.5	7	10	7	7	6.2
3.3	Evolutionary history	4.5	5	10	~~1~~	2	~~5.1~~
3.3	No opposing evidence	3	5	10	6	9	5.7
3.3	Generalisation	3	~~0~~	10	7	7	~~4.6~~
3.4	Species comparisons	4	7	10	8	7	6.5
3.4	Complex design	4	6	10	7	8	6.3

Note. All means of inquiry in this table are arguments. Unpromising arguments are marked by horizontal lines through some of their (sub)evaluation results, indicating criteria scores below bars and affected promise scores. § = section; Avail. = data availability; Cost = resource cost.

to which the argument can be applied (7), and gathering this evidence bears, in the main, a low resource cost (8). The argument from complex design has thus moderate to high promise (6.3). Both evolutionary arguments evaluated in this section are, as expected, the most promising means of inquiry for supporting consciousness as an adaptation without knowledge of its function/s.

Conclusion

To facilitate a comparison of the 10 means of inquiry evaluated in this chapter, their subevaluation and promise scores are listed in Table 3.1, and their dimensional profiles shown in Figure 3.1. The aim of all of the evaluated arguments is to support the hypothesis that consciousness is (or was) an adaptation while disregarding the potential function/s of consciousness. This category of means of inquiry has the lowest and most restricted relevance scores in the evaluation, which explains why the arguments' relevance scores vary between 3.0 and 4.5, thereby limiting their theoretical promise

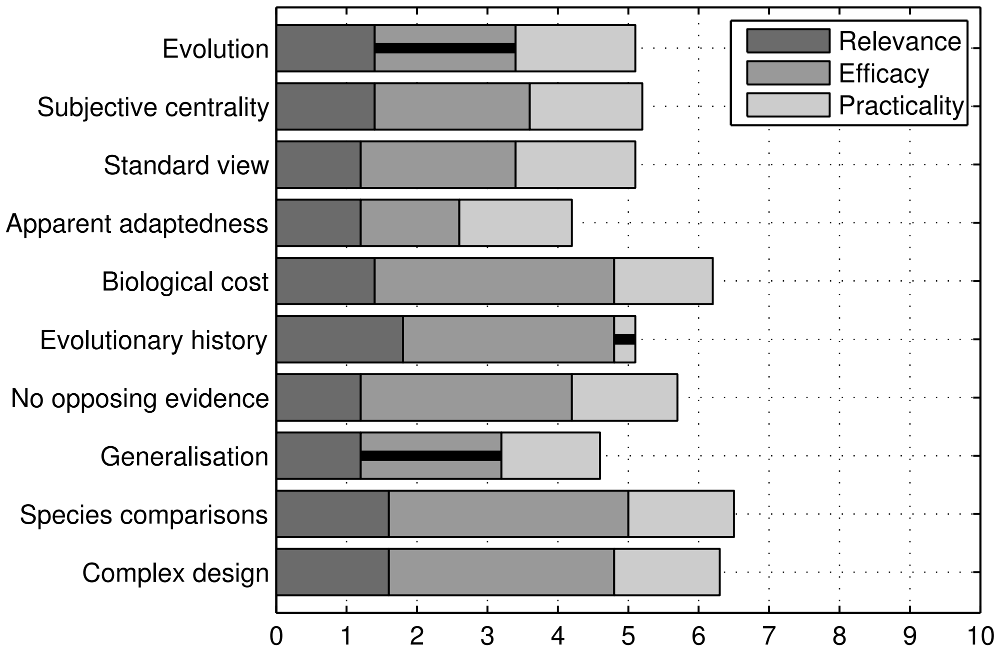

Figure 3.1. Dimensional promise profiles of the arguments evaluated in this chapter. All of these arguments were aimed at showing that consciousness is (or was) an evolutionary adaptation, hence their similar relevance scores. Unpromising arguments are marked by black horizontal bars, indicating at least one criterion score below the bar within the respective dimension.

scores to a range of 2.4 to 7.8, apart from plain *unpromising*. Three arguments are unpromising, two because of a false premise and one because of very poor data availability. The remaining arguments differ widely in their efficacy (3.5 to 8.5), but are all highly practical means of inquiry (7.0 to 8.5). The arguments that pick out features of consciousness that are directly associated with natural selection—complex design, biological cost, and distribution—are most promising for supporting consciousness as an evolutionary adaptation. These arguments should be employed on each kind of consciousness in question to find out whether their results converge.

There is an interesting discrepancy between these evaluation results and a number of the cited application examples. Most of the arguments have been found to be moderately to highly promising, and it is impressive how much evidence about consciousness as an adaptation can be mustered without any consideration of what it might be an adaptation for. Yet the cited examples tend to perform significantly worse. The most plausible explanation of this discrepancy is that the circumstances under which the arguments were employed differed markedly from the ideal circumstances assumed in the evaluation of their promise. One factor may be a poor evolutionary understanding, which, by the way, is not unique to consciousness studies (cf. R. C. Richardson, 2007; Workman & Reader, 2008). Another probable reason is the bias that I have suggested exists in many researchers in this area for consciousness being central and an adaptation. This suggestion is supported, for instance, by brief arguments being given with little or no evidence, which suffer more often from seemingly careless mistakes than fundamental problems, by the scarcity of both adaptation opponents and opposing evidence, and by the adaptation position being proclaimed the standard view. Perhaps the adaptation bias is a larger problem than suspected in section 2.1. Yet, if applied properly, these arguments are likely to make a moderate to high contribution to the research goal by helping to determine whether consciousness is an evolutionary adaptation.

The approach I have called *methodological adaptationism without hypothesised functions* is attractive for at least three reasons: It is practicable because it does not need a consensus, or even a hypothesis, on the function/s of the different kinds of consciousness. Even if we knew the function/s of consciousness, the present approach would be useful because it promises an independent analysis of consciousness, which could then be compared to the known function/s of consciousness. The approach also offers a first step toward an evolutionary explanation of consciousness. For example, the present approach could help to identify the adaptive significance of consciousness, and it can provide important information for discovering its evolution. However, some of the arguments in this section would greatly benefit from, if not require, knowledge of the function/s of consciousness (e.g., argument from apparent adaptedness or subjective centrality). Moreover, asking whether consciousness is an adaptation or not in isolation seems of limited interest; we also want to know how it evolved and what its function/s are. Common strategies for identifying the function/s of different kinds of consciousness and their evolutionary development are evaluated in the following chapter.

4 General strategies for reasoning about consciousness

"Once upon a time, there was no consciousness on Earth. There were animals, but they were aware of nothing. Their actions were instinctive responses" (Elbert, 2000, p. 288). Most statements of evolutionary or functional hypotheses about consciousness have less dramatic beginnings. Either way, we may wonder how these hypotheses were constructed and whether they merit provisional acceptance. The evaluation in this chapter therefore asks whether the strategies used to generate and develop such hypotheses may contribute to the construction of a scientifically respectable evolutionary explanation of consciousness. A major challenge for the evaluation is the fact that the majority of authors do not explain the origin of either their hypotheses or support for them, that is, if they provide any. Thus, despite the existence of many hypotheses about the function/s and evolution of consciousness, there is relatively little material on which to base the evaluation of the general strategies presumably used in their construction.

The two strategies evaluated in the first two sections of the chapter are characterised by the source of their input: introspection and common sense, respectively. However, how these types of input are treated is at least as important as their value for studying consciousness in general. The cores of the three reasoning strategies discussed in the third section of the chapter consist of the drawing of distinctions, the use of analogies, and inferences from models or theories of consciousness. In the last section of the chapter, I first examine a strategy called *contrastive analysis*, which here refers to the comparison of empirical findings about more and less conscious but otherwise equivalent processes. I then analyse the related strategy of excluding either unconsciously performed functions as potential functions of consciousness or alternative explanations of consciously performed functions. In the conclusion I compare the dimensional and promise evaluation results of these general reasoning strategies.

4.1 From introspection to functions

Concerning introspection as a point of departure for consciousness research, Ramachandran (1980) asked, given that "Galileo and Newton began with observations ... Why sneer on the same approach being used for studying our own conscious experience?" (p. 8). Whatever the details of the varied historical treatment of introspection, the common view among contemporary consciousness researchers writing about introspection is that introspective reports are the primary source of information about consciousness (e.g., Baars, 2003b; Block, 1995; Dehaene & Naccache, 2001; Haynes et al., 1998; Marcel, 2003; Overgaard, 2003). This is not surprising because *consciousness* refers to a first-person phenomenon, at least in current thought. It is thus very likely that introspective data have a role to play in the evolutionary study of consciousness. The purpose of this section is consequently to determine, based on uses and suggested uses of introspective data in the relevant research literature, what role or roles these data may fulfill in this area and with what promise.

Introspection is a vital but not infallible source of information about consciousness. Many illuminating examples and evidence of its unreliability have been given (e.g., Marcel, 2003; Schwitzgebel, 2008). There are several reasons why it has been, and may still be, called "a dangerous tool to handle carelessly" (Hebb, 1954/1994, p. 833). Among the issues already mentioned in section 2.1 are the potential inaccuracy of reports on mental processes generally (Nisbett & Wilson, 1977) as well as the potential incompleteness of reports on conscious contents, for example, due to language constraints (Schooler & Fiore, 1997). Introspective reports are affected by implicit theories that are, among other things, based on the individual's folk and scientific knowledge at the time (P. M. Churchland, 1985; Hebb, 1954/1994). A further problem is related to what Mandler (1975) termed "the uncertainty principle of psychology" (p. 239): Not only may the task influence what is reported, the conscious contents themselves may change through introspection. However, methodological improvements can increase the validity of introspective reports (e.g., Ericsson & Simon, 1980; Velmans, 1993), and such reports are widely relied on in psychological research. Besides, most of these concerns apply equally to reports about anything but one's own mind (Marcel, 1988; Velmans, 2007). Hypotheses about consciousness based on introspection are likely to be corrected and refined with scientific progress, but for now they are an important basis for such progress.

Problems with introspection, such as the ones listed above, are not uniform; that is to say, employing introspection in some research situations is more problematic than in others. How problematic depends on, for example, whether introspection is used to generate ideas, to engage the audience, or to collect data as part of a confirmatory study. It also depends on the mental aspect under investigation (e.g., sensory perception, mental imagery, reasons for actions, problem solving; Baars, 2003b) and, as already pointed out, the design of the study (e.g., instructions to report attended vs. unattended information, with or without delay, and with or without additional required inferences; Ericsson & Simon, 1980). Moreover, introspective reports can be a confounding factor in consciousness studies (Overgaard, 2004; see section 4.4). Note that introspection itself is not a single process, and the different kinds of introspection have different epistemic qualities (Butler, 2006; Prinz, 2004). One should thus recognise that introspection, like any other measurement method, has measurement limits (Baars, 2003b). The way forward for introspection as a measurement method is, in Lieberman's (1979) words, "to identify systematically the conditions under which it is most likely to prove useful" (p. 332). My interest here is not in this general project but in the determination of the usefulness of introspection specifically for contributing to evolutionary explanations of consciousness, to which I now turn.

Strategies with introspective input. So how have introspective data been used, or suggested to be used, in the evolutionary study of consciousness? I am not aware of any such suggestions where the authors have not also followed their own suggestions in the same work. The question therefore reduces to how introspective data have actually been used in this area. I begin with an example that helps to explain the intended generality of introspection uses. Bringsjord (1997) cited his own experience of not being phenomenally consciousness of a particular activity against Block's (1995) objection to an argument for phenomenal consciousness having the function of facilitating creativity. Bringsjord (1997) first avowed "since I often do what [the hypothetical character] Brown does, I can inform Block that the answer is 'No' " (p. 146). He then generalised that "*all* of us, I venture, have experienced" (p. 146) a particular type of experience. I could exclude such explicit generalisations to everyone's experience from the evaluation below, but the difference is often merely rhetorical: Introspective evidence has been variously described either in the first-person, second-person (typically in imperative mood as an activity for the reader), or third-person perspective. In all of these cases, researchers need to generalise their hypotheses or findings from introspection to other organisms at some point to be able to make a contribution to the evolutionary study of consciousness.

Usages of introspection in research on the evolution and function/s of consciousness vary widely. For instance, Block (1995) defended his use of an introspective example of phenomenal consciousness without access consciousness as part of his argument that functions of the latter kind of consciousness have been illicitly attributed to the former. Others have used introspective evidence in arguing for their philosophical frameworks of the nature and derived or supported function/s of consciousness (e.g., Tye, 1996; Van Gulick, 1997). I address the strategy of inferring functions of consciousness from models or theories in section 4.3, but not the construction of these models or theories. The strategy discussed in section 4.4, contrastive analysis, can take different types of input including introspective evidence. For example, Roth (1999) and Merker (2005) suggested evolutionary functions of consciousness based on "the contrast between the types of information that are and are not included within its compass" (p. 89). From this brief overview, it should be clear that introspection may contribute to evolutionary explanations of consciousness in many different ways. Some of these, however, are not easily evaluated as strategies beyond their particular application, and others are evaluated in later sections because their overarching strategy can also be pursued without introspective input.

The strategy I do evaluate in this section involves a more direct connection between introspection and function/s of consciousness. An early modern application of this reasoning moves from consciously experienced pleasures and pains to their evolutionary functions as seen in the accompanying behaviours. Herbert Spencer (1872) included a chapter entitled "Pleasures and Pains" in his *Principles of Psychology*, in which he stated that "pleasures are the incentives to life-supporting acts and pains the deterrents from life-destroying acts" (p. 284). Spencer's account was endorsed by James (1879) and Baldwin (1896). A few years after Spencer, Darwin (1887/1958) wrote the following comment about the evolution of "sentient beings" (p. 89) by natural selection (which was, however, not published for over 75 years; Badcock, 2000):

> An animal may be led to pursue that course of action which is the most beneficial to the species by suffering, such as pain, hunger, thirst, and fear,—or by pleasure, as in eating and drinking and in the propagation of the species, &c. (p. 89)

To come to this result, early modern researchers presumably interpreted the correlation of their own conscious pleasures and pains with their actions in the short and long term.

Far from having been superseded, the introspection–function strategy for pleasures and pains is still being used. In most instances its use leads to the same conclusion: "It is good that we feel pain. It keeps us from being burned, cut, and maimed. ... the capacity to experience pleasure and pain is a design solution that Mother Nature has often used in different lineages of locomoting organisms" (Flanagan, 1992, p. 42; see also Cairns-Smith, 1999; Pinker, 1997). Yet the same strategy has also been employed to argue against the usefulness of consciousness: "We know from experience that in fact rapid withdrawal from, say, a hot surface, typically precedes the conscious experience of heat" (Macphail, 2008, p. 99). Is this discrepancy in outcomes a problem for the strategy? Yes, in the sense that it highlights the difficulty of inferring causation from correlation, apart from showing how input dependent the strategy is. However, the quoted examples could simply differ in the periods of effectivity considered: Even if conscious pain does not initiate immediate action, which is an important insight, consciousness of pain *could* be beneficial, for instance, for overriding withdrawal reflexes or helping to avoid similar situations in the future.

Introspective evidence from many other situations has served as input to the introspection–function strategy. In order to illustrate the variety of applications of this popular strategy, I present four examples that have yielded well-known functions of consciousness. First, Mandler (1975) asserted that the trouble-shooting function of consciousness becomes apparent when originally learned but now automatically performed actions, such as typing, driving, or participating in small talk, are suddenly interrupted by a failure, such as a stuck key, a faulty brake, or an indignant direct address. Second, Baars (1997b) illustrated his list of cognitive functions of consciousness with an imagined situation meant to resemble everyday problems of our evolutionary ancestors. He introduced this self-experiment for the reader as follows: "Consider what would happen *at this very moment* if you were accosted by some large, aggressive, and dangerous beast, such as a full-sized angry bullock. How would you cope? And how would you use consciousness, as such, to survive" (p. 157)? Baars emphasised that this reference to the reader's experience (or imagination) is "a key source of evidence" (p. 157), hence not merely didactic.

Another application of the strategy is Gregory's (1996a, 1996b, 1998) comparison of the vivid visual experience when looking at something with the much dimmer experience when remembering or imagining it. He took this difference to indicate that the phenomenal qualities of conscious perceptions "serve to flag the present" (Gregory, 1996b, p. 377) so that it is not confused with the remembered past or anticipated future. This seems necessary because perceptions are based on stored knowledge to at least 90% (Gregory, 1998). Regarding my above warning not to overinterpret the perspective in which introspective evidence is described, it is worth noting that Gregory stated the experiential difference between perception and mental images from memory or imagination first in general terms (1996b), then about himself (1996a), and later as a self-experiment for the reader (1998).

The final introspective example comes from the most widely known evolutionary account of (reflective) consciousness. Humphrey's (1982) use of the introspection–function strategy begins with its description: "From all I know about myself, what strikes me—and seems to give some kind of cutting edge to consciousness—is this" (p. 474).[1] In short, he claimed that the regular accompaniment of his waking behaviour by conscious experiences had led him to consider his consciousness as identical to the neural mechanism controlling his behaviour. He thus contended: "*In so far as I am conscious*, I can see as if with an inner eye into my own" (p. 475) neural behaviour-control mechanism. Consciousness provides him with a way to make sense of his own behaviour and, through this explanatory model, of the behaviour of others. This would have been more advantageous for our ancestors living in collaborative groups than explanations based merely on input–output observations of their own and others' behaviour.

Evaluation of the introspection–function strategy. At least three reasons have been advanced against the usefulness of introspection for discovering the function/s and evolution of consciousness. The first reason has to do with animal consciousness: Corballis (2007) maintained that our introspection helps little with investigating the evolution of consciousness because "the identification of consciousness in non-human species must surely be based largely on behavioral rather than introspective evidence" (p. 572). The quoted statement seems reasonable, although, as

[1] Although *all one knows about oneself* is not limited to introspective knowledge, it appears to be Humphrey's focus in the present paper. In a 1987b lecture, he said about consciousness that because "every one of us has experienced it in our own person, to consult oneself may not be such a bad plan after all" (para. 2) and that "commonsense coupled to a bit of self-analysis suggests that consciousness is a source of information, and that this information is very likely about brain-states" (para. 37).

Corballis acknowledged, at least macaque monkeys appear capable of providing behavioural introspective evidence (e.g., Cowey & Stoerig, 1995; Myerson et al., 1981). However, difficulties with obtaining introspective reports from nonhuman animals do not constitute a problem for the strategy in question, because the strategy aims to follow a different path to a better understanding of the evolution of consciousness, namely via its evolutionary function/s in us.

The second and most commonly cited reason for the alleged unsuitability of introspection for studying the function/s of consciousness is that introspection has no access to permanently unconscious mental processes. Barlow (1980) mistrusted introspection in part because it seems to conceal that unconscious processes can guide our actions. Roth (2000a) explained that we do not have introspective access to the unconscious reasons of our conscious experiences owing to the anatomy and physiology of our brains; for example, we are not aware of the strong influence that the unconsciously operating limbic system has on our conscious experience. This information may or may not be important for figuring out the function/s of consciousness. Rosenthal (2008) argued that we cannot determine introspectively whether suggested functions of consciousness, such as rationality and intentional action, also occur unconsciously. I agree that we cannot do so directly by introspecting, but we could notice the result of such a function having been executed after a period in which no related content was present in consciousness (cf. examples above given by Mandler, 1975). Besides raising doubts about the veracity of the results of introspection, which I address shortly, the criticisms point to a serious limitation of introspection when employed in isolation for directly inferring function/s of consciousness: the inaccessibility of unconscious processes. The upside of this limitation is that it serves to demarcate consciousness from unconsciousness, a contrast that many researchers have used to suggest what consciousness might be good for.

The third reason against the usefulness of introspection is that it does not inform us directly about the survival value of consciousness. In a nutshell, "Nature does not tell us what our organs are for" (Barlow, 1980, p. 82). Barlow (1980, 1987) illustrated this principle with pain, which feels unpleasant but is protective, and love, which poets' introspective descriptions do not portray as serving to propagate the human species. Ultimate motivations do indeed not need to be conscious, so long as the associated proximate mechanisms promote them (Rossano, 2003). In line with this, William James (1890) wrote about instincts:

> It is not for the sake of their utility that they are followed, but because at the moment of following them we feel that that is the only appropriate and natural thing to do. Not one man in a billion, when taking his dinner, ever thinks of utility. He eats because the food tastes good and makes him want more.[2] (p. 386)

In order to learn more about the evolutionary function/s of consciousness, one should therefore consult the effects it has, if any, particularly on behaviour (Barlow, 1980, 1987). This is an important clarification about the introspection–function strategy, but not a common mistake in the contemporary consciousness research literature. Introspective messages are nowadays rarely, if ever, used to directly infer function/s of consciousness without additional consideration of, for instance, unconscious processes or associated behaviours.

To be able to apply the promise evaluation tool to the introspection–function strategy, the strategy's aim needs to be established. All applications of the strategy of which I am aware, including the examples described above, occur in an evolutionary context. Although the function that is derived from introspective evidence is frequently not called *evolutionary* (or comparable terms), in none of the applications is there any doubt expressed that the identified function/s of consciousness could differ from its evolutionary function/s. So the aim of the introspection–function strategy is to identify evolutionary function/s of consciousness from introspective evidence about current human consciousness. The strategy might thus help to discover survival value/s of consciousness, albeit only in humans for now; and it might provide limited additional information about the corresponding proximate mechanisms, making it moderately relevant to the research goal (5.5).

The assessment of the strategy's input should concentrate on introspective evidence, which is to be used in combination with other input, such as evidence about associated behaviours or unconscious processes. The additional input cannot easily be evaluated in general, as it differs between applications of the strategy, at times completely. However, this is only a significant impediment to the evaluation if the additional types of evidence are much weaker than introspective evidence, which is unlikely after what has been said above. Because introspective evidence might be obtained and

[2]It is likely that today more people in Western countries think of the consequences of what and how much they eat, possibly indirectly linked to evolutionary reasons; regardless, the obesity epidemic demonstrates his point.

used in various ways, it is difficult to assign a single quality score to it. As a rough estimate and under the assumption of ideal circumstances, including the collection of introspective data from a good number of people, the quality of the input is likely to be high (8). This score takes into account that the introspection–function strategy relies on a better, or more direct, criterion for consciousness than most other evaluated means of inquiry.

Some pitfalls of the strategy's procedure can be avoided by not depending on introspective evidence in isolation and by implementing the procedure carefully (e.g., not confusing introspective messages with survival values), both of which are assumed here. However, these measures cannot prevent erroneous inferences from correlation to causation and failures to detect the potential uselessness of (particular kinds of) consciousness. Another concern is that human consciousness in 21st-century civilisations might differ markedly, at least in some respects, from its biologically evolved version. Because there are several ways of using introspective and other input as part of the introspection–function strategy, and these are not specified by the strategy, another somewhat crude generalisation is required in its evaluation: Its procedure might work at least moderately well overall (6). Turning to the strategy's practicality, introspective evidence is readily available, as is pertinent additional input for the most part (9). Employing the strategy involves gathering this input and reasoning to potential function/s of consciousness; it has thus a low resource cost (8). Taken together, the promise of the introspection–function strategy for contributing to an evolutionary explanation of consciousness is high (6.7). As in the previous chapter, I report the remaining means of inquiry in this chapter and their evaluations in less detail.

4.2 Folk psychological input

What about the things we all know about consciousness? With very few exceptions, all humans live with consciousness day in, day out. It would make sense if, over many generations and in each lifetime, much knowledge about consciousness is accumulated. Such knowledge would be part of folk psychology, which helps us, as everyday people, to generate common-sense explanations and predictions of behaviour. Folk psychology can also play an important role in the work of researchers, specifically psychologists, as acknowledged in a standard behavioural research methods textbook: “Social and behavioral scientists rely heavily on commonsense notions regarding behavior, thought, and emotion” (Leary, 2004, p. 7). This suggests that folk psychology could be useful in the investigation of consciousness too, and “it is exactly in this area that the appeal to common sense has always seemed most cogent” (Hebb, 1954/1994, p. 831). Consequently, I consider its usages and their potential to assist with evolutionary explanations of consciousness in this section.

Usages of common-sense knowledge. Folk psychology is commonly used to argue that consciousness is functional or, more precisely, that consciousness is not functionless or epiphenomenal. Such a use of common sense could well be the basis of the following statement by Roger Sperry: “It is just that I find it difficult to believe that the sensations and other subjective experiences *per se* serve no function” (as cited in Voneida, 1998, p. 1078). Similarly, Humphrey (1987b) stated that he found definitions of consciousness as a useless phenomenon “counter-intuitive and wholly unconvincing” (para. 26). And Gregory (1996a) supported the assumption that phenomenal properties of conscious experience are causal by commenting that “this fits common sense” (p. 756). Because these uses of common sense are aimed at showing that consciousness has a function, I do not evaluate them here (as stipulated in section 2.4 for means of inquiry in the a-function category). Related arguments are based on the inconceivability either of consciousness not having a function (e.g., James, 1890/1907) or of it having a function (e.g., Macphail, 2000, who based his whole account of the evolution of consciousness on this). Yet how close a hypothesis is to the truth is not indicated by how imaginable (P. S. Churchland, 1996) or intelligible (Wright, 2007) the hypothesis is. In any case, I do not assess these uses of common sense either, because their aim is as little evolutionary as the folk-psychological arguments for or against epiphenomenalism of the metaphysical or causal-role variety.

Appeals to common sense have also been made in arguments for consciousness being an evolutionary adaptation. For example, Bringsjord et al. (2002) considered possible answers to the question, designated as $\mathrm{Q1}_{\mathrm{P}}^{\mathrm{Z}}$, of why phenomenally conscious humans have evolved instead of functionally identical unconscious zombies. They used a good dose of common sense to dismiss the answer, designated as A1, that consciousness does not aid survival and hence that there is no particular reason for conscious beings having evolved instead:

> A1 is really not an answer to $\mathrm{Q1}_{\mathrm{P}}^{\mathrm{Z}}$; and as such it's profoundly unsatisfying

> (if the informal poll we've taken is any indication). It even seems downright bizarre to hold that the phenomenon that makes life worth living (Wouldn't you be depressed upon hearing that starting five minutes from now you would have the inner life of a slab of granite?) is a fluke. (p. 121)

This is a nice example of reliance on common sense: To reject the possibility that consciousness is not an adaptation, the authors depend exclusively on their own and others', including the presumed readers', dissatisfaction with and opinion of the suggestion as too out of the ordinary.

It is important for the evaluation of the use of folk-psychological input to determine the most relevant subgoal for which it has been employed. The categories that are distinctly more relevant than the a-function and an-adaptation categories have the subgoals of identifying the evolutionary function/s of consciousness and identifying its evolutionary origin and history. A contender for the most relevant category is a reason that Humphrey (1995) gave against his earlier account of the evolution of reflective consciousness outlined in the previous section. The reason has to do with reactions to a consequence of his theory, namely that most animals and babies are not conscious: "I couldn't sell this idea even to myself, let alone to my nonphilosophical friends" (p. 200). This reason, as well as others, did not motivate him to abandon his previous theory but to realise its proper domain and to construct a theory of conscious sensation, which I mentioned in section 3.3. In this particular case, common-sense beliefs seem to have triggered, and thus contributed to, evolutionary theorising about consciousness. However, it appears more like a fortuitous coincidence than planned use of a means of inquiry and is therefore not suitable for the present evaluation.

The most relevant evaluable applications of common-sense knowledge in consciousness research use it as support for evolutionary function/s of consciousness. For instance, it is likely that some of today's uses of the introspection–function strategy for pleasures and pains, which was evaluated in the previous section, are based on folk psychology to a significant degree; the same goes for conscious intentions to act (cf. Libet et al., 1983) and free will (cf. Wegner, 2002). Clear expressions of the use of folk psychology in this area are difficult to find, which is why I return to Humphrey's (1987b) writing for another valuable example. He introduced his evolutionary function of reflective consciousness with the example of a dentist enquiring whether his patient was feeling pain and then noted "Common sense tells me that when a person describes his states of mind, either to me or to himself . . . , he is making a revealing self-report" (para. 27). Humphrey's theory does not follow folk psychology in all respects, but he evidently employed folk psychology to support his hypothesis that consciousness is useful because it is informative.

Evaluation of folk input. The use of common-sense knowledge in research has received much bad press, both in general and in consciousness studies. The main problem is that folk psychology and intuitions derived from it can turn out to be (probably) wrong about the function/s of consciousness (e.g., Kinsbourne, 2006; Rosenthal, 2008), but they can also be (probably) right (e.g., Barlow, 1987; Goldman, 1993; Morsella, 2003). Of course, this is true for most sources of information, so maybe the criticism, especially when nonphilosophical, is that folk psychology is wrong more often (but see Pinker, 1997, 2006). Graham and Horgan (2002) advocated a charitable approach to the folk psychology of the function/s of phenomenal consciousness in saying that "folk psychology need not be completely *correct* . . . , but its homely platitudes are innocent until proven guilty" (p. 65). I agree that common-sense knowledge can be a very useful starting point. However, both folk psychology and the intuitions it inspires change, and these changes by and large lag behind increases in scientific understanding, which expectedly applies to consciousness as well (Banks, 1993; Lamme, 2006). We should thus be careful not to let common-sense ideas about consciousness hinder scientific progress. As addressed briefly in section 1.1, consciousness itself is a folk-psychological term that will likely be elaborated through research (including its division into components; Frith & Rees, 2007). An interesting question is where folk-psychological beliefs about consciousness come from, but because I am not aware of any suggested or actual attempts to tackle this in connection with the evolution or function/s of consciousness, I do not deal with it here.

The use of common-sense beliefs should be evaluated here, as already established, as a strategy in the what-adaptation category, that is, for identifying the evolutionary function/s of consciousness. This strategy's moderate relevance to the research goal (6) is slightly higher than that of the introspection–function strategy, for two main reasons: Folk psychology extends to nonhuman animals (think of the conscious pleasure and pain examples rather than the reflective consciousness one) and might tap into how people actually use consciousness or what role/s it plays in their lives. The quality of folk-psychological input to the strategy is moderate (5) due to the many potentially distorting influences, such as individual (e.g., psychological, biological), social, political, and economic interests. Folk psychology is simply less interested in the truth than is scientific psychology; its goal is rather to allow the best possible explanations and predictions of the behaviour of people in relevant social groups based on user-friendly rules of thumb.

If the present strategy consisted of importing folk-psychological accounts of the function/s of consciousness into scientific research as definite statements about its evolutionary function/s, the strategy's procedure would be poor. However, Humphrey (1987b) did not trust common sense blindly. For example, he expected his suggestion that the evolutionary function of consciousness is to provide descriptions of brain states to be perceived by most people as odd and unconvincing. The procedure of such less naive uses of folk psychology could work moderately (4), the problem now being that the strategy does not specify how to deal with the common-sense input, other than not to believe it unquestioningly. There also seems to be no consideration of the facts that folk-psychological function/s of consciousness may not be identical to its evolutionary function/s and that scientific knowledge may already have advanced past folk-psychological knowledge.

Regarding the practicality of the strategy, the availability of—rather general—folk-psychological input might be high (7); this depends on the reference group (probably not "most people in the world [who] probably don't even know they've got a brain," Humphrey, 1987b, para. 29) and how important consciousness has been within folk psychology. The resource cost of using this strategy is low (8), which may still be higher than expected, but researchers implementing this strategy need to consult with lay people. If used carefully, folk-psychological knowledge could thus make a moderate contribution to the research goal (5.7). Having considered two common sources of information about consciousness—introspection and folk psychology—that supply input to, on the whole, poorly defined procedures, my focus now shifts to the procedural cores of three general reasoning strategies.

4.3 Distinctions, analogies, and models

In his influential 1975 paper, Mandler stated that "the important advances in our excursions into consciousness must come through the usual interplay of empirical investigation and imaginative theory" (p. 231). There certainly is a wealth of what could be regarded as the early beginnings of such theories, as described in section 2.1.

There are at least a few dozen hypotheses about the evolution or function/s of consciousness in the research literature that have been provided with very little evidence, most of which are referred to as *speculations* by their authors. Even if support is given for such hypotheses, the means of inquiry used to construct them often remain unclear. Out of close to 100 works with distinguishable reasoning steps that lead to or support a hypothesis on the evolution or function/s of consciousness, I have selected three reasoning strategies to be evaluated in this section. They are less specific than an individual applications' reasoning steps and also than the arguments in the last chapter, but more specific than groupings of means of inquiry such as reasoning or providing empirical evidence.

Drawing distinctions. A good example of the first strategy is Dretske's (1997) drawing of three distinctions in order to reveal the evolutionary functions of consciousness. He explained his strategy as follows:

> It seems to me that the flurry of interest in the function of consciousness betrays a confusion about several quite elementary distinctions. Once the distinctions are in place—and there is nothing especially arcane or tricky about them—the advantages (and, therefore, the good) of consciousness is obvious. (p. 2)

He distinguished creature consciousness (without it the creature is "a vegetable," p. 5) from state consciousness, conscious states either as being made conscious by the creature becoming aware of them (no function) or as making a creature conscious (which is then their function), and the associated conscious knowledge from conscious experience (maybe for the identification and recognition of objects). To give another example, Cairns-Smith (1999) constructed a "phylogeny of qualia" (p. 277) based on his classification of phenomenal consciousness as raw perceptual, interpretative, intellectual, coercive, volitional, or background feelings. Relatedly, Mesulam (1998) used dimensions of human consciousness, such as intentionality, introspection, and theory of mind (other introspecting selves), as hypothesised evolutionary milestones.

Can the evaluation tool determine how promising this general strategy is for contributing to an evolutionary explanation of consciousness? I think so, although the little information available on the strategy increases the uncertainty in the assigned scores. As the latter two examples indicate, distinctions have been drawn with the aim of identifying stages in the evolutionary history of consciousness. If the distinctions are well-supported, the relevance of this strategy to the research goal is at least moderate (5.5): It may help to discover major stages in the evolution of consciousness and suggest evolutionarily relevant differences between them. The input to the strategy could include comparative and developmental evidence, yet in the examples consists solely of knowledge of contemporary adult human consciousness of unspecified origin. Note that the relevance score takes this limited applicability into account. Provided it is scientific knowledge from which the distinctions are drawn, the strategy's input could be of very high quality (9). The procedure might work moderately well (5), the main concern being that the evolutionarily relevant distinctions and their order may not be obvious from the input and mistakes not noticed (historical data could be useful with this). Data about adult human consciousness is highly available (7), though it may not have afforded suitable, sophisticated distinctions so far (but see Leary & Buttermore, 2003, , on the evolution of self-consciousness in different domains). The evidence-based drawing of distinctions has a low resource cost (8). The calculated promise score for drawing distinctions is moderate (6.5).

Analogical reasoning. The second strategy in this section was used by Nichols and Grantham (2000) as part of their complexity argument. As outlined in more detail in section 3.4, these authors argued that because the structure of phenomenal consciousness is analogous to that of the lateral line system in fish, which is an evolutionary adaptation, phenomenal consciousness also is an adaptation. Climbing the relevance ladder, Baars (1998) promoted a theatre metaphor for consciousness with implications for the evolutionary function/s of consciousness:

> Many proposals about brain organization and consciousness reflect a single underlying theme that can be labeled the 'theater metaphor'. In these views the overall function of consciousness is to provide very widespread access to unconscious brain regions. Such access is needed for global activation, co-ordination and control. (p. 62)

Finally, Tannenbaum (2001) suggested that consciousness might be usefully viewed as emerging from the organising mechanism of a biological sense of consciousness that provides adaptively useful information about the brain. It is then the case that "much of what we know about the evolution and role of the conventionally recognized senses should help us understand the evolution and role of the sense of consciousness, and of consciousness itself" (p. 377). Tannenbaum proposed that the sense of consciousness evolved, like other complex senses, through the integration of existing systems. Furthermore, it "might have acquired the capacity to scan and organize its environment into useful patterns, just as some other senses acquired the capacity to scan and to organize their environments" (p. 380). The sense of consciousness would thereby have increased the knowledge that the organism can have about itself and its environment, as all other biological senses did. It differs from other senses, however, in its complexity, comprehensiveness, and unique effects, which include our conscious experience.

Analogical reasoning has thus been used for each of the three evolutionary subgoals specified in the relevance dimension of the evaluation tool. The third application example above suggests that, in the most relevant category, the use of analogies could contribute a potential, or even likely, partial evolutionary history of consciousness with adaptive significance/s, albeit in very broad strokes (6.5). The input to this strategy comprises knowledge about consciousness that enables the selection of an analogy and knowledge about the analogous thing, so that their relevant similarities and dissimilarities can be identified. The quality of this input could be high under ideal circumstances (8). Depending on the relevance and number of similarities, among other things, the strategy's procedure may work moderately in the evolutionary study of consciousness (6).

The availability of input to the strategy that enables it to be used successfully is difficult to assess. The analogical thing needs to be suitably comparable to consciousness and have known developmental stages for the strategy to work. Although consciousness can surely be compared to many things, it is doubtful that there are good analogies for each kind of consciousness readily available that allow us to learn more about those kinds' potential evolution (4). (Better input would be available if the aim was only to identify the survival value/s of consciousness, but the associated increase in promise would be offset by a lower relevance score, viz. 5.5.) Employing this strategy requires the identification of a suitable analogy, which has very variable resource costs, and then inferences from its comparison with consciousness, so might carry low resource costs on average (8). The combination of these subevaluations suggests that, if suitable evolutionary analogies can be found, analogical reasoning in this context is highly promising (6.6).

Inferring from models. The third general strategy in this section involves inferences from models or theories of consciousness and therefore not reasoning from evolutionary theory. For example, Shallice (1972) identified consciousness with the selector input to the dominant action system in an information-processing model, based on similar properties. As a consequence, consciousness inherits the dual functions of the selector input, namely strongly activating the action system and setting its goal. Computational models can also help to study the function/s of consciousness (A. Seth, 2009), but the focus here is on the use of the model strategy for subgoals that are more relevant to an evolutionary explanation of consciousness. Barlow (1980) derived an evolutionary function of consciousness from his theory of its nature: "The view that consciousness arises in interpersonal relations . . . gives a clear answer to the question about its functional role and survival value" (p. 88), that is, to induce humans to communicate with each other. And lastly, Grossenbacher (2001) inferred from his neural theory of consciousness, the *access mediation model*, how the contents of consciousness may have changed during evolution. Although content types can become less conscious over evolutionary time, consciousness has probably evolved to include an increasing number of different content types. In addition, Grossenbacher explained why the most recent content types may be the most salient ones in consciousness.

Even if function/s of consciousness are what is most commonly inferred from models or theories of consciousness, this strategy has also been used to infer its evolutionary function/s and partial evolutionary history. Because additional details are likely to be obtained, such as about the functioning of consciousness and interactions with other phenomena, the strategy has a high potential to contribute to the research goal (7). In each specific case, the quality of the input to the strategy is a function of the quality of the chosen model or theory of consciousness. However, the quality of this type of input is high (8), though not higher because models and theories necessarily make simplifying assumptions. If a suitable model or theory is chosen, the information of interest may follow more or less directly from it without much room for error. At other times, more reasoning steps may be necessary, and these are not specified by the strategy. Nevertheless, the procedure can be expected to work at least well under ideal circumstances (8).

The critical issue for employing the model-inference strategy is not its resource cost, which is very low (9), but the availability within a decade of suitable models or theories of the different kinds of consciousness from which their evolutionary functions and histories could be inferred. There are models and theories of consciousness, but so far they allow rather limited evolutionary inferences at most (4). In sum, the promise of model inferences for contributing to the research goal is high (7.3). Having covered three possible reasoning strategies that may play a role in constructing theories, I now return to strategies that work much more closely with empirical evidence to adequately cover both aspects of the required "interplay of empirical investigation and imaginative theory" (Mandler, 1975, p. 231).

4.4 Contrastive analysis and exclusion

"In *The Principles of Psychology* (1890/1983), [William] James suggests a way of focusing on the issue of consciousness by *contrasting* comparable conscious and unconscious events" (Baars, 1988, p. xvi). James did not pursue this strategy himself (because he thought that psychology should only study conscious processes; Baars, 1988), but this is how Baars (1988) introduced the term *contrastive analysis* into consciousness studies. He later explained that "this 'method of contrastive analysis' is a generalization of the experimental method, with consciousness as a [independent] variable" (Baars, 1994, Abstract, para. 3). This extremely popular strategy has been applied in many different areas:

1. A popular everyday example is an experienced car driver driving on a familiar route and holding a conversation at the same time; when an unexpected hazard presents itself, the driver's focal attention switches back to the traffic situation.

2. Psychological experiments have compared, for instance, conscious and unconscious stimuli, conscious and unconscious elements in memory, and new and habituated events.

3. Functional brain imaging studies have looked at brain activity, for example, during conscious and unconscious perception in paradigms such as visual masking, change blindness, and binocular rivalry.

4. Healthy brains can be compared to damaged brains or to neurological syndromes such as epilepsy, blindsight, and synaesthesia.

5. Artificial intelligence researchers have compared human capabilities with what robots can or cannot do.

This list indicates that there is no lack of input to contrastive analyses, at least some of which may have been aimed at answering functional or evolutionary questions about consciousness.

So for which subgoals of giving an evolutionary explanation of consciousness, if any, have contrastive analyses been conducted? The strategy has been used for studying the function/s of consciousness (Baars & McGovern, 1996; Frith & Rees, 2007), but not explicitly for studying its evolutionary history, as far as I know. However, contrastive analyses have frequently been used to identify the adaptive significance/s of consciousness. In section 4.1 I already mentioned examples of comparisons of the contents inside and outside of consciousness that were used to identify potential evolutionary functions of consciousness (i.e., Merker, 2005; Roth, 1999). Searle (1992) inferred such functions from case reports of behaviour during absence seizures:

> Normal, human, conscious behavior has a degree of flexibility and creativity that is absent from the Penfield cases of the unconscious driver and the unconscious pianist. ... The hypothesis I am suggesting then is that one of the evolutionary advantages conferred on us by consciousness is the much greater flexibility, sensitivity, and creativity we derive from being conscious. (pp. 108–109)

Bringsjord et al. (2002) agreed that phenomenal consciousness enables creative cognition and gave the additional reason that researchers have made several attempts to engineer artificial creativity without any success, another example of the contrastive strategy.

Evaluation of contrastive analysis. Contrastive analysis is generally seen as a valuable strategy in consciousness research. For example, Velmans (1994) commented that "one could hardly take issue with the usefulness of contrastive analysis, as the method (Hume's 'method of difference') is as old as empirical science" (para. 1.2). And Mangan (1995) asserted that, "as a pragmatic method for organizing experimental results, contrastive analysis is extremely useful" (para. 1.1). It is, of course, not without its problems, some of which have been pointed out in the literature. Block (1995) called attention to the importance of properly identifying the presence and absence of relevant kinds of consciousness in the compared conditions or situations, in order to avoid ascribing the function of one kind of consciousness to another kind. Other criticisms have highlighted likely confounds, such as performance (Lau, 2008), unconscious processes, attention, and introspective reports (Overgaard, 2004). These difficulties threaten the conclusions reached by contrastive analysis, yet they can be bypassed, for instance, by following the cited critics' recommendations and suggested solutions for dealing with these confounds.

Contrastive analyses that are aimed at identifying evolutionary function/s of consciousness are likely to contribute some additional details but probably little contextual information, making them moderately relevant to the research goal (6). There is no restriction on the type of input other than that it should allow a comparison between more conscious and less conscious conditions or situations. Hence, the input should allow the strategy to work very well under ideal circumstances (9). The contrastive approach is generally sound, but there are likely confounds and no clear links to biological evolution (7): Most suggested or performed comparisons simply assume that the identified function/s have evolutionary significance. Given the thousands of relevant published studies (Baars, 2003b), the accessibility of either original data or research findings as input to contrastive analyses is very high (9), though this does not apply to all kinds of consciousness equally. If new data are collected for comparison, the resource cost is moderate; if results of existing studies are compared, it is low. Here I compromise between the two by assigning their average score to the present strategy (7). Overall, then, contrastive analysis is a highly promising strategy (7.2), although its evolutionary connection tends to be weak.

Excluding functions. A strategy that is even more concerned with unconscious processes than contrastive analysis is exclusion. This strategy seeks to narrow down the potential functions of consciousness by showing that particular candidate functions are in fact performed unconsciously. (Authors who exclude potential functions of consciousness may not have this motivation, yet this does not invalidate the evaluation of this strategy in the present context.) An example of exclusion is Morsella's (2005) conclusion about certain informational conflicts: "At a minimum, these phenomena demonstrate that conscious processing is unnecessary to integrate information from sources as diverse as different modalities. Intermodal cross-talk can occur without it" (p. 1003). It was taken for granted at the time of Descartes that consciousness enabled voluntary action and rational thought; however, the latter has been shown to occur unconsciously (Frith & Rees, 2007), as has the former. For instance, Lau (2009) presented evidence against several alleged evolutionary functions of consciousness associated with voluntary action. Exclusion thus appears to be part of the progression from a folk theory of consciousness to its scientific equivalent.

The exclusion strategy is best evaluated as being aimed at excluding potential current evolutionary functions of consciousness (i.e., what-adaptation/s category). The strategy's contribution to the research goal depends on the strength of the discounted candidate function/s. The knowledge gain from exclusion is minimal if the function does not occur in the organism in question (say, photosynthesis in humans) and maximal for contemporary strong contender functions, such as in the last two examples above. Even when the knowledge gain is maximal, the strategy is unlikely to provide additional details about consciousness and its function/s or evolution. Consequently, it receives the lowest possible score in its category (5), indicating that it is moderately relevant to achieving an evolutionary explanation of consciousness. Like contrastive analysis, the present strategy allows any type of input, which is of very high quality in a best case scenario (9). The exclusion strategy does not prescribe how the dissociation between consciousness and the function/s is to be established (the inference is unfortunately often from inessentialism to epiphenomenalism; see section 1.4). In the absence of such instructions, the strategy underperforms slightly but nevertheless works well (8). Much input is available for the strategy to be implemented (9) and at a low research cost (9). Based on these subevaluations, the exclusion strategy is as promising for increasing our evolutionary understanding of consciousness as contrastive analyses are (7.2), despite its opposite aim.

Conclusion

To make the comparison between the means of inquiry evaluated in this chapter easier, their summarised subevaluation scores are presented in Table 4.1 and their dimensional profiles in Figure 4.1. All means of inquiry were evaluated for the research subgoal of identifying the evolutionary function/s of consciousness, except the means in the third section of the chapter, which were evaluated for additionally identifying aspects of the evolutionary history of consciousness. Correspondingly, the relevance scores of these seven general reasoning strategies vary from moderate to high (5–7). The efficacy of all strategies is high (7–8.5), with the exception of the only moderately efficacious use of folk-psychological input. The procedures of some of the other strategies also work only moderately well, but the quality of the input they all take is high to very high. The practicality of all means of inquiry is also high to very

Table 4.1
Subevaluations of General Means of Inquiry for Reasoning about the Evolution and Function/s of Consciousness

			Efficacy		Practicality		
§	Means of inquiry	Relevance	Input	Procedure	Avail.	Cost	Promise
4.1	Introspection	5.5	8	6	9	8	6.7
4.2	Folk psychology	6	5	4	7	8	5.7
4.3	Distinctions	5.5	9	5	7	8	6.5
4.3	Analogies	6.5	8	6	4	8	6.6
4.3	Models or theories	7	8	8	4	9	7.3
4.4	Contrastive analysis	6	9	7	9	7	7.2
4.4	Exclusion	5	9	8	9	9	7.2

Note. All means of inquiry in this table are strategies. § = section; Avail. = data availability; Cost = resource cost.

high (7.5–9), apart from that of analogical reasoning and inferences from models, which are both limited by a moderate availability of input. Although the evaluated means of inquiry are very general reasoning strategies with different weaknesses, they are all moderately to highly promising for contributing to evolutionary knowledge about consciousness (5.7–7.3).

What the means of inquiry in this chapter have in common is that they are general reasoning strategies that make their contribution to the present research goal to a significant degree by relying on the content and quality of the input they receive. This means that the evaluated strategies benefit from a wide applicability, but are very input dependent. For some of these strategies, the modest availability of input calls into question their current practicality. In all cases but one was the quality of input evaluated as high to very high. This is because the strategies are capable of taking high-quality input, and assuming that they do helps to determine their potential or promise. However, the high input scores to some extent mask the fact that the quality of input is much lower in most of the examples cited in this chapter. (Note that I have intentionally omitted examples of lower quality.) The challenge for using such general means of inquiry, then, is to obtain suitable high-quality input. In the next chapter, I examine somewhat more specific and clearly more evolutionary

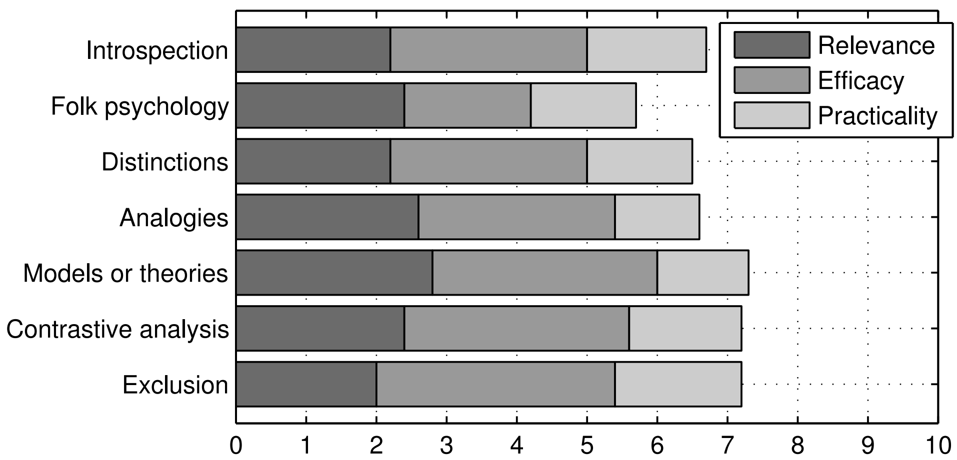

Figure 4.1. Dimensional promise profiles of the seven general strategies for reasoning about consciousness evaluated in this chapter.

means of inquiry, which also depend on input but do (a bit) more of the work toward the research goal themselves.

5 Evolutionary means of inquiry into consciousness

What we really want when investigating the biological evolution of consciousness is one or more ways of looking back at the origin of consciousness and its subsequent development. Unfortunately, we cannot record the corresponding historical data at the time of occurrence, nor do we know of any direct fossil evidence left by early consciousness, nor has anyone so far seriously suggested recreating some of the relevant conditions in a laboratory. As already pointed out, however, these difficulties are not unique to the evolutionary study of consciousness (Polger, 2007; Polger & Flanagan, 1999), and the associated methodological questions may be the most foundational in evolutionary behavioural sciences generally (Gangestad & Simpson, 2007a). Researchers in these sciences have developed several methodological approaches for studying past evolutionary changes to mind and behaviour and their effects. This chapter evaluates the promise of those evolutionary means of inquiry that have been used or suggested to be used in consciousness studies.

The first section of the chapter deals with the *natural method*, a means of inquiry that prescribes the integration of evidence about consciousness from different disciplines. The natural method could have been classified as a general strategy and placed in the previous chapter; however, it has been recommended by its author for studying the evolution and function/s of consciousness and is meant to include evidence from evolutionary biology. In the following section, I analyse a means of inquiry that is commonly employed in evolutionary biology and evolutionary psychology, namely the reverse engineering of organisms' traits. I then discuss how much we may hope to learn about the evolutionary development of consciousness from its ontogenetic development. The last section, about the comparative method, considers comparisons of the behaviour and brains of animals both across species and time. I summarise the evaluation results of these evolutionary means of inquiry in the conclusion.

5.1 The natural method and associates

"The question is by what method consciousness is to be studied. I propose that we try the most natural strategy, what I call the *natural method*, to see if it can be made to work" (Flanagan, 1992, p. 11). With this statement Flanagan (1992) introduced the natural method. Despite his promotion of it since, it is not widely cited. Other authors, however, agree with his claim that "*the natural method* ... will ... yield success in understanding consciousness if anything will" (Flanagan, 1995a, p. 1104). For example, Dooremalen (2003) deemed the natural method "the way to find answers to questions about consciousness" (p. 7) and was "convinced that if we can solve the problem of phenomenal consciousness [its realisation and evolutionary function] this method will do the job" (p. 70). So how promising is the natural method for contributing to a scientifically respectable evolutionary explanation of consciousness? And why has it not been cited more?

The natural method is a methodological strategy that gives equal consideration to any kind of evidence about consciousness. The three key sources of information about consciousness within the natural method are phenomenology, psychology, and neuroscience. Flanagan (1992) explained:

> The object of the natural method is to see whether and to what extent the three stories can be rendered coherent, meshed, and brought into reflective equilibrium. The only rule is to treat all three—the phenomenology, the psychology, and the neuroscience—with respect. (p. 11)

It is to be expected that the different kinds of evidence will be weighed differently in the end; the point is not to privilege any kind of information prematurely (Polger & Flanagan, 1999). Right from the start Flanagan included evolutionary biology as an additional potentially relevant source of information about consciousness and asserted the usefulness of the natural method for constructing a theory of consciousness covering its origins and functions.

To illustrate the natural method in action, I summarise Dooremalen's (2003) only explicitly evolutionary case study. Dooremalen purportedly combined phenomenological, psychological, and biological evidence to explain the experience of facial beauty. He took patterns in the requests for plastic surgery to suggest that there are intersubjective and intercultural standards of beauty. He then cited psychological studies indicating, first, that averaged faces are judged more attractive than actual faces and, second, what features further increase the perceived beauty of faces. These facial features have been found to be markers of high fertility, good genes, and good health. Dooremalen concluded that "these feelings [of beauty evoked by faces] are nature's way of telling us that this [potential] mate is a good candidate for reproduction" (p. 155) and, more generally, that "phenomenal experiences ... provide us (and other animals) with a fast way to pick up highly complex information in a glimpse" (pp. 156-157). This by itself does not show that the conscious experience of beauty affects behaviour in an evolutionarily relevant way, but the case study does show the coordination of different kinds of evidence according to the natural method.

One may wonder what is new about this multidisciplinary strategy. A coevolution of phenomenological, psychological, and neuroscientific research has been repeatedly advocated, albeit under different names and with variations (e.g., P. S. Churchland, 1986; Revonsuo, 2006; Varela, 1996; Wimsatt, 1976). This is not surprising, for example, in light of O'Keefe's (1985) claim that a complete theory of consciousness should comprise a psychological model linked to knowledge from introspection, neuroscience, behavioural sciences, evolutionary biology, and sociology. Polger and Flanagan (1999) admitted that consciousness researchers already do—at least implicitly, to some extent, and some of the time—seek to balance the different kinds of evidence. In fact, consciousness researchers nowadays often explicitly integrate different kinds of information about consciousness, for example, in behavioural-cum-neuroimaging studies of participants reporting conscious contents (cf. natural-method examples in Flanagan, 1992, pp. 13-19), and this might help explain why the natural method is not more widely cited. However, Flanagan (1992) asserted that employing the natural method was not standard practice, and Polger and Flanagan complained that "almost everyone who has approached questions about consciousness has done so by privileging one source of information above the others" (p. 243). An unwarranted preference could certainly hinder research progress, and so an important feature of the natural method is considering all relevant informational sources on an equal footing.

Evaluation of the natural method. The research subgoal in relation to which the natural method is evaluated here is identifying the evolutionary function/s of consciousness (what-adaptation/s category; see section 2.4). Although the natural method has also been recommended for learning more about the evolutionary history of consciousness (Flanagan, 1992, 1995a), making it a contender for the origin-and-history category with the highest relevance to the research goal, its adequate evaluation in this category would require more information than is contained in the available very general strategy and applicability statements. Apart from applications of the natural method in support of nonadaptation claims (discussed later in this section), Dooremalen's (2003) case study above is the only acknowledged evolutionary application of the natural method I know; all other acknowledged instances of its use are about different properties of consciousness. Dooremalen's case study thus facilitates the evaluation of the natural method as a means of inquiry for identifying the survival value/s of consciousness.

So if the natural method is employed with the aim of identifying evolutionary function/s of consciousness, its potential contribution to the research goal is high (6.5), as it is likely to provide additional relevant information about consciousness. Lacking specific instructions on how to apply the natural method with this aim, I assume that it is unlikely to contribute knowledge of (differing) past evolutionary function/s and detailed contextual conditions. The natural method might well be capable of doing so, but so far this has not been clearly suggested. Flanagan and Polger (1995) may have come closest to such a suggestion when they called for the cooperation of "biologists, cognitive scientists, and philosophers of mind with interest and expertise in evolutionary theory ... plus comparative psychologists, ethnologists, paleontologists, zoologists and neuroscientists, among others" (p. 321) for solving the problem of how consciousness evolved, yet they did not mention the natural method. By excluding, for the purpose of this evaluation, past function/s and contextual conditions as likely results of the natural method, a lower demand is placed on the data required as input.

The natural method takes as input evidence about consciousness from all relevant sources, whether neuroscience or folk medical practice (Polger & Flanagan, 1999); it stipulates that all of these sources must be treated with equal respect, independent of their quality. Note that the quality of this input is to be evaluated in relation to the efficacy of the natural method and that this efficacy is not impaired by some lower-quality evidence, because such evidence is simply weighed less heavily when following the procedure of the natural method. Naturally, including and assessing all relevant kinds of evidence is very desirable for a means of inquiry (10). Moreover, cross-disciplinary integration of evidence is very important when studying multidisciplinary problems, such as the adaptation status of psychological traits (Schmitt & Pilcher, 2004) or the evolution of consciousness. However, without a more specific articulation of the procedure of the natural method, there are doubts about how cogently and exclusively the conclusions follow from the input (7).

Regarding the availability of input to the natural method, Polger and Flanagan (1999) commented that "what one gets, automatically, as it were, by taking up ... the natural method, is an immense body of data" (p. 235; as for the contrastive analysis evaluated in section 4.4). The natural method can indeed take any available relevant evidence as input, but evidence about the evolutionary function/s of consciousness is not as abundant as evidence about some of its other properties (and evidence about its past function/s and contextual conditions is much less so). As Dooremalen's (2003) case study illustrates, relevant evidence, though available, may not be sufficient for drawing a convincing conclusion about the survival value/s of consciousness. Because the available evidence clearly allows the natural method to be used but not to reach its potential, the data availability in the present case is moderate overall (6). Employing the natural method involves locating, assessing, and coordinating relevant evidence from various disciplines, even if known by other means to be inadequate, and thus carries a moderate resource cost (6). The synthesis of all subevaluation scores indicates that the natural method is highly promising (7.2).

Deflationary evolutionary explanations. It is likely that at least some kinds of consciousness are not adaptations and hence do not have any evolutionary function/s. The natural method can be used in such cases also, leading to *deflationary evolutionary explanations* (Polger, 2007; the term *deflationary* was introduced in this context by Flanagan, 1999/2000, and Polger & Flanagan, 1999). Given that this type of application of the natural method has its own name, the question arises whether it requires a separate evaluation. We need to know, for example, whether the natural method has been used, or suggested to be used, merely for supporting the hypothesis that a particular kind of consciousness has no evolutionary function/s (an-adaptation category) or for developing a more detailed alternative account of that kind of consciousness as a nonadaptation (what-adaptation category). This is where two further evolutionary applications of the natural method come in handy.

The first of the two acknowledged uses of the natural method for giving deflationary evolutionary explanations is Flanagan's (1995a, 1995b, 1999/2000) argument that dreaming is not an evolutionary adaptation. In brief, ponto-geniculo-occipital waves during REM sleep activate visual areas (and memories), which, when combined with the adaptive tendency of the cortex "to make sense out of experience"

(Flanagan, 1995b, p. 25), largely accounts for phenomenological reports of dreaming. Flanagan claimed that the activation of visual areas serves their early development but has no other obvious developmental function. Even once in place, he argued, dreaming is unlikely to have been affected by positive natural selection. Furthermore, the phenomenology of dreams is not in agreement with predictions of their content based on potential memory-related functions of REM sleep. The integration of this phenomenological evidence with the aforementioned neuroscientific data and evolutionary reasoning led Flanagan to hypothesise that dreaming is a by-product of both the possession of waking consciousness and adaptive brain processes during sleep.

The second deflationary evolutionary explanation, given by Sufka (2000; Sufka & Turner, 2005), is of chronic pain. Although acute and persistent pain are adaptive, as they assist with the prevention of further injury and with recovery, chronic pain has no known use, only disabling effects. Sufka (2000) suggested that the key to solving the problem of "how something so maladaptive to an organism as chronic pain could have ever possibly evolved in the first place" (p. 156) is the striking similarity between the neurochemical mechanisms involved in chronic pain, specifically wind-up, and those subserving learning and memory, specifically long-term potentiation, a form of activity-dependent neural plasticity. Chronic pain may thus result from neural plasticity in the pain sensory system. And despite being maladaptive, it could not have been selected against because it is tied to subchronic pain on the one hand and neural plasticity on the other, both of which are phylogenetically old and highly adaptive. Sufka and Turner (2005) concluded that susceptibility to chronic pain appears to be a maladaptive by-product of pain and neural plasticity.

The interest here is not in the promise of these two deflationary evolutionary explanations, but in that of the deflationary application of the natural method. What these two examples demonstrate is that the natural method has been used in attempts to explain certain kinds of consciousness as by-products of specified adaptations, rather than merely to support the view that these kinds of consciousness are not adaptations. Essentially the same strategy is used whether the natural method is employed to give an adaptation account or a deflationary evolutionary explanation. Although the aims in those two situations are opposite with respect to the adaptation status of consciousness, both types of application of the natural method aim to discover the reasons for the postulated status (what-adaptation category). Because of these equivalent aims and the same strategy (and equally little information for accurately estimating performance on the other criteria), the evaluation of the natural method need not be repeated in writing for deflationary evolutionary explanations.[1] Deflationary applications of the natural method are therefore roughly as promising (7.2) as its applications to alleged adaptation kinds of consciousness.

5.2 Reverse engineering consciousness

In biological evolution every variation of an organism that arises is tested against existing alternatives in the relevant context. Depending on, among other factors, the relative effects of this variation on the organism's survival and reproduction, the variation is retained or not. Virtually everything biological that we see today is the cumulative result of many such generational cycles as well as the basis for future developments. Thus, when studying a trait of an organism at the present time, "the biologist is in the position of an archaeologist who uncovers a machine without any written record and attempts to reconstruct not only its operation but also its purpose" (Lewontin, 1978, p. 164). Such reverse engineering in biology has been promoted by Dennett (1995a) as "extremely fruitful and, in fact, unavoidable" (p. 213). This view has been extended explicitly to psychology, for example, by Pinker (1997), who contended that "psychology is engineering in reverse" (p. 21). Taken together, these considerations suggest that the application of reverse engineering to consciousness is a worthwhile target of evaluation.

Application examples. William James (1879) already engaged in reverse engineering consciousness, as is apparent in his summing up of the core of his argument:

> *A priori* analysis of both brain and conscious action shows us that if the latter were efficacious it would, by its selective emphasis, make amends for the indeterminateness of the former; whilst the study *à posteriori* of the *distribution* of consciousness shows it to be exactly such as we might expect in an organ added for the sake of steering a nervous system grown too complex to regulate itself. (p. 18)

James characterised consciousness and its selection function at different levels, identified a potentially corresponding adaptive problem, and compared the distribution

[1] If the chronic pain account were the only evolutionary application of the natural method so far, the natural method would score slightly higher on relevance (7) owing to a more historical and contextual perspective, but lower on general data availability (5), resulting in almost the same promise score (7.3).

of consciousness in animals to that hypothesised for an optimally designed trait with the same function. Note the similarity of his wording to that of evolutionary biologist Williams (1966): "Suppose we did find some features of the feeding activities of earthworms that were ... exactly what we should expect of a system designed for soil improvement. We would then be forced to recognize the system as" (pp. 18–19) an evolutionary adaptation.

A well-structured and clear application of this general strategy to consciousness is Pinker's (1997, 2004) reverse-engineering analysis of access consciousness. He first explained that any information processor's access to information must be limited because information (including information processing) is costly in terms of space, time, and energy. A computational system should therefore be designed so that only currently relevant information is accessible to the processor in question. Pinker (1997) claimed that this design specification comprises "the engineering specs of information access, and thus the selection pressures that probably gave rise to it" (p. 137). Next he compared his independent optimality analysis with the empirical features of access consciousness in us. According to Pinker's (2004) strategy, if the criteria derived from the design specification for the relevance of information match the empirical features of access consciousness, we may conclude that it is an adaptation.

Four empirical features of access consciousness are pertinent to the design specification because they each prioritise relevant information or discard irrelevant information (Pinker, 1997, 2004):

1. The perceptual contents of access consciousness are at intermediate levels of mental representation, between sensory input and abstract cognition. This makes sense for an information processing system because computations at lower levels are based on rules that are mostly constant over one's lifetime and also context-independent to a certain degree, whereas higher levels lack important information.

2. Access consciousness has a serial, limited-capacity focus corresponding to attention and working memory. The spotlight of attention helps to detect conjunctions of features. Because conjunctions are combinatorial, unconscious parallel processing cannot be designed within the space of a brain to be capable of detecting all possible conjunctions. Relatedly, slow serial memory retrieval conforms to many principles of an optimal information-retrieval system, aimed at

retrieving the currently most useful information. Examples of such principles are retrieval by frequency and recency.

3. Conscious states have emotional colourings. These activate goal states, which in turn initiate operations for reducing discrepancies with present states. Sources of emotional colourings such as pleasures and pains are closely related to survival and reproduction, these sources being precisely those factors that could not be taken for granted in our evolutionary environment, such as food (but not oxygen), romantic partners, and surviving children.

4. Conscious experience includes the feeling of being in control, of making decisions about what actions to take. An executive control mechanism has the benefit of facilitating the selection and coordination of a single course of action for a single body.

Pinker's (1997, 2004) descriptions of these features of access consciousness are based on everyday experiences and research studies. In sum, Pinker gave independent engineering reasons for the existence of access consciousness and four of its features in order to support its explanation as an adaptation.

Because applications of reverse engineering to consciousness differ, I briefly summarise four additional examples. Dessalles (2001) drew an analogy between, on the one hand, the experiential spaces associated with each sensory modality, which are characterised by continuous, directional, and regular stimulus-experience projections, and a large dynamic range, and on the other hand, optimal projections in communication theory with the same four properties. He concluded that phenomenal consciousness has been optimally designed by natural selection for perceptual discrimination. Merker (2005) argued that the systematic absence from consciousness of "sensory preliminaries and motor sequels of central control functions" (p. 95) is a design feature: Mobile animals with centralised brains and diverse complex sensory and motor systems would have benefited from decision making within a single stable summary world. Baars (1993) explained the biological cost of a single limited-capacity stream of consciousness with the benefit of global workspace architectures, which can, by combining the activities of various specialised processors, solve problems that specialised processors cannot solve by themselves. Making use of Nichols and Grantham's (2000) *principle of adaptive complexity*, according to which "the features that make an organ complex likely contribute to its adaptedness" (p. 667), also amounts to reverse engineering: A complex trait's structure may assist in figuring out what the trait was designed to do. For instance, phenomenal consciousness is unified and has multiple input mechanisms, which suggests that it integrates information (like the similarly structured lateral line system in fish; Nichols & Grantham, 2000). In each of these examples, an aspect of consciousness has been reverse-engineered, often with the help of an analogy, in search of that aspect's evolutionary function/s.

Evaluation of reverse engineering. Criticisms of reverse engineering offer clues as to how well and for what this strategy might work. The trait to be reverse-engineered is at the outset assumed to be an evolutionary adaptation, in line with methodological adaptationism. To be able to implement the reverse-engineering strategy, one then needs a hint of what the trait might have been designed to do (Pinker, 1997). Even if the trait is an adaptation and we have an idea of what its evolutionary function/s might be, there are many reasons why its design may not be optimal (see section 1.2). What is more, the trait's function/s might be underdetermined by its structure (R. C. Richardson, 2007), simply because "the same feature may serve different ends ... [just as] the same end may be served by different features" (Griffiths, 1996, p. 524). Further, R. C. Richardson (2001) complained that reverse-engineering analyses with a priori unknown design constraints are compatible with different historical routes, and Griffiths (2001) that a reverse-engineering explanation can be highly plausible, correctly predicting the observed trait, and the best one available, yet false (e.g., as shown by comparative data). Against the expectations underlying these criticisms,[2] a single means of inquiry is seldom sufficient to support a broad hypothesis beyond reasonable doubt, and the focus of reverse-engineering analyses is on the adaptive advantages of traits, not their precise historical development. A benefit of the ahistoricity of reverse engineering is that it allows the development of hypotheses about functions based on more readily available evidence.

Some problems with reverse engineering have also been mentioned in papers on consciousness. For example, Sufka and Turner (2005) explained chronic pain as a maladaptive by-product using what they called a *bottom-up research strategy* and commented that "anyone who tried to 'reverse engineer' chronic pain ... would go badly astray" (p. 255) because it was not designed to solve an adaptive problem.

[2]These criticisms served to argue against an over-reliance on reverse engineering in evolutionary psychology, narrowly construed, and for the use of the comparative method.

This is the above point about the adaptation assumption of reverse engineering applied to consciousness: Attempts at reverse-engineering a kind of consciousness are bound to fail if that kind of consciousness is not an adaptation—neither of which may be obvious. Relatedly, Welshon (in press) set aside reverse engineering of consciousness because he believed that many current human traits have not been selected for their current adaptiveness, which would invalidate inferences about their evolutionary functions. So when considering an aspect of consciousness that might have been affected by recent changes, the existence and nature of these changes should be investigated first (Sleutels, 2008), that is, here, before attempting to reverse-engineer the corresponding aspect of consciousness.

The application of the promise evaluation tool to reverse engineering requires its aim to be established first. The application examples above make evident that reverse engineering is employed in consciousness research with the aim of identifying evolutionary function/s of consciousness. Two inferences are involved in achieving this aim: "*from* structure *to* function and then *from* function *to* history" (R. C. Richardson, 2007, p. 44). The latter inference is where the analogy of the present strategy with the reverse engineering of technology or software breaks down. When reverse-engineering such a product, the interest is normally in the functioning and purpose of the product and its components, not in its historical development. However, when reverse-engineering a biological trait, the more contentious historical inference is essential for determining its naturally selected function/s.

The relevance of reverse engineering to the research goal is moderate (6): Because of its aim, this strategy belongs in the what-adaptation/s category, and its use can reasonably be expected to provide some additional knowledge, such as knowledge about important characteristics of consciousness, its functioning, and possible selection pressures. The principal input to reverse-engineering analyses consists of empirical data on consciousness, for instance, from laboratory experiments or introspection. If data on consciousness in today's animals are used and their sources are appropriately taken into account, which is suggested, for example, by Pinker's (2004) discussion of the conscious feeling of being in control, the data quality is likely to be very high (9). This is also true for other reverse-engineering input, such as hypotheses about functions, suitable analogies, and the bases of optimality analyses, again under ideal circumstances.

Most of the criticisms of reverse engineering concern its procedure, which I rate as moderate overall (5). The strategy does connect input to output and is commonsensical, but there are problems with its inferences and assumptions. For example, the structure-to-function inference can go wrong because of underdetermination, and the function-to-history inference blindly assumes that the trait in question is an adaptation largely unaffected by culture. Other questionable assumptions are that the trait solved an adaptive problem optimally and that trait and problem can be analysed in isolation. The availability of the data that have been described above as input to reverse-engineering analyses is high (8). Employing this strategy requires the gathering of such mostly available empirical data and their analysis, possibly including an independent optimality analysis, and thus carries a relatively low resource cost (7). Despite concerns with its procedure, reverse engineering is on the whole a highly promising strategy (6.7)—if executed well—for studying the evolution and function/s of consciousness.

5.3 Parallels between ontogeny and phylogeny

A completely different approach to the investigation of the evolution of consciousness was suggested by Baldwin (1896), based on his assertion that "in the life history of the organism we have the problem of development actually in a measure solved before us" (p. 302). Given the difficulties with obtaining evidence about the evolutionary history of consciousness, data on its development in present-day organisms could be a handy resource indeed. The corresponding strategy consists of examining how consciousness develops in the individual organism and generalising the findings to the biological evolution of consciousness. This strategy may not only help to learn more about the evolutionary stages of consciousness: Baldwin claimed that "the biologist no less than the psychologist must ... resort to this field of investigation [children's mental development] if he would finally settle the function of consciousness in evolution" (p. 302). Although this strategy has been employed little in consciousness research compared to the other evolutionary means of inquiry evaluated in this chapter, Baldwin's statements indicate that its promise is well worth evaluating.

A crucial question concerns the intended closeness of the parallels between the ontogeny and the phylogeny of consciousness. If ontogeny were thought to recapitulate phylogeny strictly, investigations of individual development could be taken to provide direct knowledge about the evolutionary history of consciousness. However, the parallels between ontogeny and phylogeny are most likely weaker (see next subsection). This has been acknowledged by researchers writing about the evolution of consciousness. For example, Baldwin (1906) devoted half of his first chapter to the discussion of systematic deviations from strict parallelism in mental development. Herrick (1945) stated that the "developmental series [of experience] shows many similarities with the phylogenetic series ... though the parallelism is by no means close" (p. 63). And Reber (1992) pointed out that "the recapitulation is not 'literal'" (p. 111) and that "a gentle form of recapitulation holds" (p. 123) for the emergence of consciousness. Having clarified that the present strategy assumes limited ontogeny–phylogeny parallelism only, I first consider example applications of this strategy and then examples of other uses of ontogenetic findings in evolutionary studies of consciousness.

Example uses of ontogeny. As called for above, Baldwin (1896) generalised from observations of the early development of human children that consciousness has the evolutionary function of enabling the acquisition of noninherited things: "The instinctive equipment of the lower animals is replaced by the plasticity necessary for learning by consciousness" (p. 301). He summarised further results of his application of the ontogeny–phylogeny strategy, but did not demonstrate the strategy in action. A better source of parallelism examples is Baldwin's (1906) book *Mental Development in the Child and the Race*. Having recommended the use of the present strategy in research on the evolutionary history of consciousness, he outlined four evolutionary stages of mental development, corresponding approximately to invertebrates, lower vertebrates, higher vertebrates, and humans, and identified these stages in the development of human infants. Conversely, he described four stages of human children's consciousness of other persons, namely others as objects, others with personality, the child as subject, and others as subjects, the first three of which he claimed to have clear parallels in other animals. These are early illustrations of stages in human individual development found to resemble evolutionary stages.

In two other cases in which the ontogeny–phylogeny strategy has been used, the evolutionary development is presented before that of the individual. Reber (1992) stated as an axiom that consciousness is evolutionarily recent; after all, "we have consciousness, protozoa do not" (p. 113). Among the predictions that he derived from this axiom and confirmed mainly by literature review is that conscious processing develops later in infancy than unconscious processing. Herrick (1945) described conscious experience in humans but believed we had no means of determining its emergence in either ontogeny or phylogeny, which is why he focused on its early precursors. He distinguished two primordial types of experience in all animals: sensorimotor and integrative experience (the latter is associated with the active preservation of the individual and the coordination of its parts), which are forerunners of human extro- and introspection, respectively. During evolution animals have tended to become more specialised to achieve a better fit with their environments in an increasing number of ways, with higher-ranked animals surviving and living better. Making use of parallelism, "the evolutionary series of progressively amplified experience may profitably be compared with the simpler and more accessible series of changes observed in the growth of the individual from egg to maturity" (p. 63). Herrick cited developmental studies of the behaviour and anatomy of many different species including humans in support of the proposed types of experience, of which the integrative type seems to be primary and develop first. Ontogenetic findings can thus support old and suggest new phylogenetic hypotheses.

More commonly, usages of alleged ontogeny–phylogeny parallels in the development of consciousness are sketchier or have different purposes. For example, Thomas Huxley (1874/1882) argued for consciousness in animals and therefore against the sudden emergence in humans of what he deemed a complex phenomenon: "We know, that, in the individual man, consciousness grows from a dim glimmer to its full light, whether we consider the infant advancing in years, or the adult emerging from slumber and swoon" (p. 233). Gärdenfors (2008) used behavioural evidence from human children and nonhuman animals to support his ordering of five levels of intersubjectivity, by which he referred to the awareness of others' consciousness. When considering how the integration of the components of consciousness develops, Lagercrantz (2008) suggested that sensory signals in the brains of human newborns are not yet fully integrated. To support his statement that "maybe there is an evolutionary parallel" (p. 171), he gave an animal example from a theoretical chapter. Finally, the only obvious use of ontogeny in Roth's (2000b) chapter "The Evolution and Ontogeny of Consciousness" occurred in his comparison of the capabilities of human children and nonhuman apes to learn theory of mind and language. These examples show awareness of the potential usefulness of parallels between the ontogeny and phylogeny of consciousness—the question is how far their use can take us toward an evolutionary understanding of consciousness.

As an aside, the concept of recapitulation already existed well before Darwin's

(1859) *Origin of Species* (Mayr, 1994), and possibly also before its discovery by the German Naturphilosophen in the 1790s. Blaise Pascal expressed the basic idea in 1647, albeit not about different species: “The same thing happens in the succession of men in general as in the different ages of a single individual man” (as cited in Martindale, 1977–1978, p. 262). It is therefore not surprising that early, culturally oriented theories of the evolution of consciousness already tended to subscribe to the recapitulation idea and rely on observations of children as support among other sources (reviewed by Martindale, 1977–1978). More recent theories of the history of consciousness that maintain that only humans with a conceptual language are conscious also draw on evidence from developmental psychology (e.g., Jaynes, 1986; Macphail, 1998, 2008). Overall, then, we might expect the recapitulation heuristic to have a certain value, but what about its biological foundation?

Evaluation of ontogenetic parallels. There may be no consensus on recapitulation in biology, but the general view seems to be that parallelism between ontogeny and phylogeny does hold to some extent, even if the original law of recapitulation has been largely discredited (M. K. Richardson & Keuck, 2002). In short, parallelism holds when new stages are added to the end of ontogenetic development, but not when new characters are introduced earlier or old ones modified or when reproduction occurs at an earlier stage (Gould, 1977; Ridley, 2004). Groups of species may differ in their susceptibility to the different types of changes (Mayr, 1994). Recapitulations have been almost exclusively found in the transformations of single characters, not whole stages as originally proposed (M. K. Richardson & Keuck, 2002). Futuyma (2005) concluded that the recapitulation “law is certainly not an infallible guide to phylogenetic history. However, embryological similarities provided Darwin with some of his most important evidence of evolution, and they continue to shed important light on how characteristics have been transformed during evolution” (p. 56). Because there is no one-to-one correspondence between ontogeny and phylogeny, care is needed when using the former to help reconstruct the latter.

The aim of employing ontogeny–phylogeny parallels in consciousness research is to learn more about the evolutionary development of consciousness. This means that the strategy fits well into the origin-and-history category of the promise evaluation tool. The relevance of this strategy to the research goal is high (7), though it has a slight bias for the history as opposed to the origin of consciousness. Its strength lies in the potential discovery of evolutionary changes of consciousness and their sequence. Additionally, parallels with ontogeny could suggest factors responsible for these changes and possibly also evolutionary function/s of consciousness. The input to the strategy consists primarily of findings on the consciousness of young human children. Good data may be difficult to obtain in some cases (see below), but even then the strategy's input can be strong (8) if the results of several means of inquiry agree.

Although the ontogenetic-parallels strategy can lead to valid conclusions, its procedure is threatened by the variable reliability of the link between ontogeny and phylogeny (6). Not only do ontogenetic changes occur relatively faster or slower, stages in the development of consciousness may be missing, additional, or in the wrong order. The strategy should include a mechanism for trying to ascertain whether recapitulation is likely to apply to the case at hand. One could, for instance, study the ontogenetic development of consciousness in other species, as Herrick (1945) did for its precursors. Or one could investigate whether the bases of consciousness are part of the main lines in the embryogenetic development of vertebrate nervous systems that Nicolau et al. (2000) asserted recapitulate phylogeny. In any case, the availability of data that allows the application of the strategy is high (7), in spite of some methodological and ethical difficulties with pre- and postnatally collecting data on consciousness. Unlike historical data on consciousness, human children are not rare, and results of developmental studies are available (see Lagercrantz, 2008; Zelazo, Gao, & Todd, 2007). The resource cost of locating these studies and collating their results would be low (7), though further studies will need to be conducted. Taken together, the promise of the ontogenetic-parallels strategy, which offers a unique access to historical developments of consciousness, is high (7.0).

5.4 The comparative method: Round two

"The comparative method ... was Darwin's favoured technique. And, since his day, it has been used sooner or later by almost all right-thinking evolutionary biologists" (Harvey & Pagel, 1991, p. 5). The comparative method is also the most popular evolutionary means of inquiry among consciousness researchers. A basic version of it was already evaluated in section 3.4 as an argument for consciousness being an adaptation. Comparisons of animal species have, in general, been used at many different levels, for example, to identify adaptations that have evolved independently in similar selective environments (analogies) and evolutionarily relevant variations of a trait in species with shared ancestry (homologies) as well as to determine when a trait has evolved. Hence, in order to use the comparative method properly, the phylogenetic relationships of examined species must be taken into account (Freeman & Herron, 2007; Futuyma, 2005; Harvey & Pagel, 1991). I therefore exclude from my analysis of the comparative method publications that discuss animal consciousness without any reference to phylogeny. The majority of consciousness researchers who have employed the comparative method have compared both animal behaviour and brains. However, these types of evidence are often not well integrated, they have been employed individually, and they face different criticisms, which is why I evaluate their use in the comparative method separately.

Today's behaviours. A memorable example of a behavioural comparison of mammals with reptiles led Sjölander (1995) to identify what he considered "one of the most important breakthroughs in the evolution of mind" (p. 5). Sjölander (1997) compared, amongst other behaviours and species, the mouse-related behaviours of cats and snakes. When a cat hears a mouse, it might search for it by sight and smell, anticipate its path behind an obstacle, wait for its emergence from a hole for many hours, and perhaps even dream about it. A snake, on the other hand, relies purely on its eyes (or heat-sensitive organs) to strike a prey, on smell to track the prey to where the venom immobilises it, and on touch to swallow the prey by its head. The snake does not anticipate the path of its prey and, even when holding the prey with its body, searches for it by smell (Sjölander, 1995). The snake's behaviour thus appears to indicate that it does not have a single centralised mental representation of reality as we and probably other mammals do. Rather, it "seems to live in several different worlds, where a mouse is not an *object* with different characteristics, but rather many 'things'" (Sjölander, 1997, p. 597). The behaviour of birds indicates that their inner worlds may be placed somewhere between those of mammals and reptiles; they either live in a single reality or in many sophisticated modality-specific ones (cf. Sjölander, 1999). These behavioural comparisons suggest a fundamental change in the experience of animals in the mammalian lineage after its divergence from reptiles, and possibly analogous experience in birds.

Because comparisons of animal behaviours differ in many respects, a couple of remarks on the strategy and two further examples are helpful. To begin with, the comparative method does not specify how behaviours should be selected for comparison, so I do not address criteria for animal consciousness here (see section 3.4 for some comments on this). Sometimes the search in comparisons is for a critical difference between human behaviour or cognition and that of nonhuman animals regarded as unconscious (e.g., Macphail, 1998, 2008), but more often it is for several species groupings according to prespecified behavioural or cognitive abilities (e.g., Gärdenfors, 2008; Oakley, 1985). Pursuing the latter approach, Roth (2001) listed nine cognitive-behavioural criteria for human consciousness, of which the first one may be found in primates and other mammals, whereas the last one is found in humans only: task-structure and tool learning by imitation, perspective-taking in deception, anticipation of future events, comprehension of mechanisms, theory of mind, mirror self-recognition, teaching, simple syntactical language, and complex syntactical language. He used evidence of the presence or absence of these criteria in mammals to identify four evolutionary steps in their cognitive abilities including consciousness. Focusing on the evolutionary origin of consciousness instead of evolutionary changes in consciousness, Cabanac et al. (2009) concluded, based on experimental evidence on emotion, sensory pleasure, play, REM sleep, and other indicators in several vertebrate species, that consciousness emerged in early amniotes approximately 325 million years ago (see Figure 5.1). These examples illustrate well enough for evaluation purposes how behaviour comparisons have been employed in different ways to learn more about the evolution of consciousness.

The aim of such behaviour comparisons is to discover the origin and development of consciousness by linking its distribution in animals to their phylogenetic relationships. The potential contribution of these comparisons to the research goal comes closest yet to the best evolutionary history of consciousness that we can currently hope for (8). Because much evidence is missing, we cannot expect, for example, to establish the responsible evolutionary mechanisms in every case or to give complete adaptation explanations. However, this strategy's contribution is likely to include evolutionary relations between different levels of consciousness in many different species, some of whose selective environments may be partially known. A potentially significant drawback of behaviour comparisons is that they can only be applied to aspects of consciousness that have clearly and consistently observable effects on behaviour. Yet unlike the argument from species comparisons in section 3.4, the comparative method can be used to investigate the emergence of distinctly human aspects of consciousness when historical evidence, such as archaeological artefacts, is taken into account (e.g., Leakey, 1994; Leary & Buttermore, 2003; Mithen, 1999; Welshon, in press). The

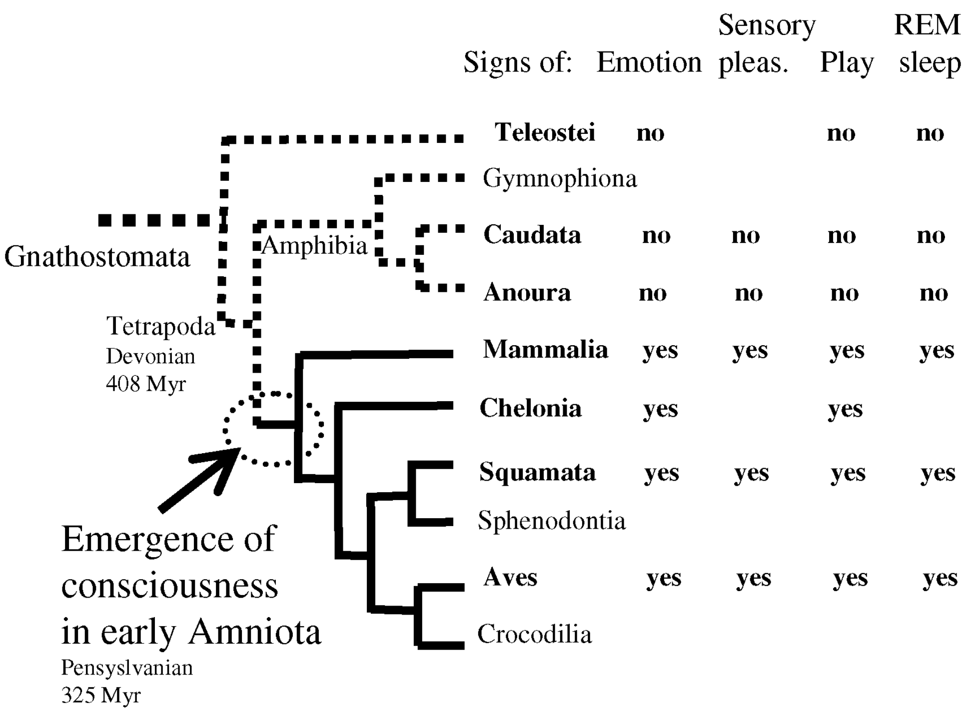

Figure 5.1. Phylogenetic tree with behavioural signs of consciousness present or absent in living vertebrates. (REM sleep signs are probably based on behaviour and neurophysiology.) Myr = million years; Sensory pleas. = sensory pleasure. From "The Emergence of Consciousness in Phylogeny," by M. Cabanac, A. J. Cabanac, and A. Parent, 2009, *Behavioural Brain Research*, *198*, p. 271. Copyright 2009 by Elsevier. Reprinted with permission.

input to this strategy is thus made up of behavioural studies of living animals (e.g., experiments, laboratory and field observations) or, less commonly, historical evidence (e.g., artefacts, skulls, bones, fossils). The high quality of well-conducted animal studies (8) may be approached through historical evidence about the human lineage under ideal circumstances.

Behaviour comparisons utilise the pattern of presence and absence of consciousness in animals well, and their procedure is generally sound (8), though the determination of animal consciousness as part of this strategy can be problematic. For one thing, the assumption that the behaviour of today's animals is the same as that of their early ancestors may well not hold. For another, it is questionable if criteria for consciousness, whether derived from theory or based on the human model, can be applied successfully to a large variety of species. To conclude from the *absence* of a criterion behaviour that the species is fully unconscious, one would have to exclude failures to elicit the behaviour and consciousness being expressed differently. To conclude from the *presence* of a criterion behaviour that the species is fully conscious, one would have to establish that the observed behaviour is not produced by a consciousness-unrelated mechanism (strict necessity is unlikely for biological traits; Polger, 2007; A. K. Seth, 2009). Examples of apparently complex behaviours that are better explained as instincts evolved for specific purposes are the communication of bees by waggle dance and the anticipation of dragon-fly larvae for their prey to reappear on the other side of an obstacle (Sjölander, 1997). The difficulties with behavioural criteria differ with the generality of the criteria (e.g., mirror self-recognition vs. behavioural complexity) and increase with the phylogenetic distance between species. My purpose here is not to question criteria for consciousness, on which all evaluated means of inquiry have to rely, only to point to potential difficulties in their generalisation to distantly related species. Most behaviour comparisons mitigate these problems by the simultaneous use of multiple criteria.

The availability of relevant data depends a lot on the particular aspect of consciousness under investigation, which influences the choice of criterion behaviours, animal species, and type of evidence. Much evidence is available from fields such as ethology and anthropology, so there should be sufficient initial input for the strategy to be pursued (7). Published behaviour comparisons on the evolution of consciousness support this assessment, yet more specific data on the chosen criteria for consciousness will need to be collected and may not be available within a decade. The main steps in using this strategy are locating and analysing evidence on criterion behaviours of potentially many different species and relating the evidence to the corresponding phylogenetic information. Because the strategy makes use of published results in the first instance, its resource cost is low (7). Altogether, the comparison of animal behaviours is highly promising for increasing knowledge about the evolution of consciousness (7.8).

Brain evolution. Comparisons of those features of animals' brains that are thought to be associated with consciousness are at least as popular for answering evolutionary questions as are behaviour comparisons, and they resemble each other in many respects. Here I concentrate mainly on their differences, starting with a brief look at some applications of the comparative method to human and nonhuman brains. A basic introductory example is Thomas Huxley's (1874) argument that lower vertebrates have a less developed equivalent of the human organ of consciousness and therefore most likely a proportionally lower level of consciousness, "which, more or less distinctly, foreshadows our own" (Huxley, 1874/1882, p. 233). He believed the anterior division of human and, by analogy, nonhuman brains to be the organ of consciousness. What is being compared in applications of the comparative method has changed and expanded since Huxley's time—a trend I expect to continue with the neuroscience of consciousness having become a very active research area—yet the strategy of brain comparisons in this area has essentially remained the same.

Which brain features are compared depends on the aspect of consciousness investigated, the approach taken, and often also on the theory of consciousness entertained. For example, A. K. Seth et al. (2005) identified three empirically established indicators of human consciousness in mammal brains, namely the presence of fast, irregular, low-amplitude electrical activity, cortex and thalamus, and widespread activity.[3] D. B. Edelman, Baars, and Seth (2005) considered related criteria for primary consciousness in birds and octopuses, based on G. M. Edelman's (1989) biological theory of consciousness. Grossenbacher (2001) inferred that consciousness has not evolved recently in humans from evidence that the cingulate cortex, hypothesised to be the central consciousness structure in his *access mediation model*, has an ancient cytoarchitecture and similar connectivity in other mammals. Taking a different approach, Roth (2001) searched for brain features among tetrapods that would correlate with the complexity of previously established cognitive abilities including consciousness and reflect "the factual or alleged superiority of humans regarding cognition and consciousness" (p. 575). After considering features such as overall brain organisation and the absolute and relative size of the brain, cortex, and prefrontal cortex (or homologues), he concluded that the best candidate feature is the number of cortical neurons and synapses, along with the brain's maturation period and Broca's area. Unsurprisingly, the structural and functional features in brain comparisons are virtually always premised on human consciousness as a model.

Despite having the same aim as behaviour comparisons and sharing many of their properties, brain comparisons are more relevant to the research goal (9). This is because they can be applied to aspects of consciousness that are apparent in behaviour and to those that are not, which is attractive if we are not just looking for adaptation

[3]This paper is primarily concerned with criteria for consciousness for use in comparative studies.

explanations of consciousness. Brain comparisons are in principle applicable to all aspects of consciousness that are generated by neural processes (in interaction with their environment), but whether this will be practicable, and if so when, is too difficult to estimate at this point. The evidential strength of brain data is also higher (9) than that of the input to behaviour comparisons. Brain comparisons have been defended as a good starting point for comparative studies because the corresponding criteria are more easily testable (particularly for homologues structures; A. K. Seth et al., 2005) and lead to more concrete results (Århem, Lindahl, Manger, & Butler, 2008). Nevertheless, assigning a score of 10 is not justified, as the input to brain comparisons may be about functional equivalents in distantly related species or about historical brains.

The other subevaluation scores of brain comparisons are the same as those for behaviour comparisons (8–7–7), even if some of the reasons differ. Regarding the procedure of brain comparisons, the search for structures that are associated with consciousness in humans risks missing evidence of consciousness associated with other structures, especially in distantly related species. That is, similar-looking homologous structures might be identified but not analogous structures. One should therefore look for functional equivalents of human consciousness indicators (e.g., D. B. Edelman et al., 2005). Moreover, the presence of a (homologous or analogous) structure that is associated with consciousness in humans does not mean that the structure has the same functionality in other species. Until we have something approaching a unique marker for a particular aspect of consciousness, it is important to use several indicators at once, as the authors of most published brain comparisons have done. The reasons given for the scores assigned to the availability of data for and the resource cost of behaviour comparisons transfer to behaviour comparisons. In sum, then, using the comparative method on the brains of different species is a highly promising strategy for investigating the evolution of consciousness (8.4).

Applications of the comparative method can be made even more promising by integrating evidence on the brains and behaviour of animals, thus pursuing a *synthetic approach* to the study of animal consciousness (D. B. Edelman et al., 2005; D. B. Edelman & Seth, 2009). It is encouraging that the majority of researchers who have utilized the comparative method in research on consciousness have attempted to do so, even if findings from the different criteria often do not agree well. A nice example of an integration of historical behaviour and brain data is Welshon's (2008) use

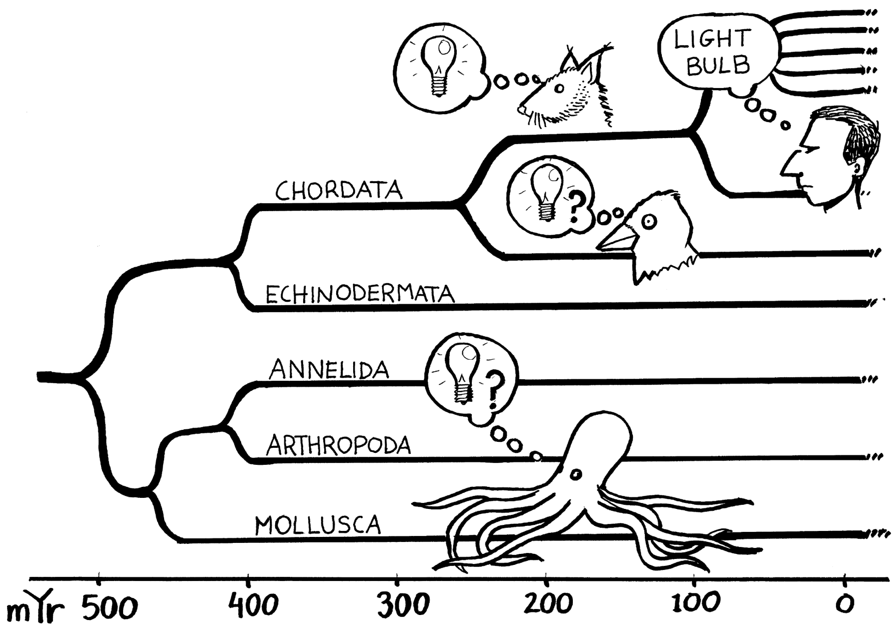

Figure 5.2. Phylogeny of consciousness based on a synthetic approach. Behavioural, anatomical, and physiological data suggest that primary consciousness emerged independently in the mammalian, avian, and possibly cephalopod lineages. Higher-order consciousness is present in modern humans. mYr = million years. From "Identifying Hallmarks of Consciousness in Non-Mammalian Species," by D. B. Edelman, B. J. Baars, and A. K. Seth, 2005, *Consciousness and Cognition*, *14*, p. 174. Copyright 2005 by Elsevier. Reprinted with permission.

of archaeological evidence (viz., of componential tool construction) and paleoneurological evidence (viz., of differences between humans and monkeys in specific parietal gyri and associated connectivity and of nonallometric parietal and frontopolar expansion in the human lineage). He argued that the implicated enhancement of human working memory in the last 1.5 million years resulted in changes to access consciousness. The complexity of brain and behaviour comparisons, including inconsistencies between them, is apparent in Århem et al.'s (2008) discussion of four basic scenarios of the evolutionary origin of consciousness: emergence in humans, in mammals, in mammals and birds independently, or in reptiles. According to D. B. Edelman's (2009) comparative analyses, primary consciousness may even have evolved at least three times (see Figure 5.2): in mammals, birds, and cephalopods.

Table 5.1
Subevaluations of Evolutionary Means of Inquiry for Use on Consciousness

			Efficacy		Practicality		
§	Means of inquiry	Relevance	Input	Procedure	Avail.	Cost	Promise
5.1	Natural method	6.5	10	7	6	6	7.2
5.1	Deflationary method	〃	〃	〃	〃	〃	〃
5.2	Reverse engineering	6	9	5	8	7	6.7
5.3	Ontogenetic parallels	7	8	6	7	7	7.0
5.4	Comparing behaviour	8	8	8	7	7	7.8
5.4	Comparing brains	9	9	8	7	7	8.4

Note. All means of inquiry in this table are strategies. § = section; Avail. = data availability; Cost = resource cost.

Conclusion

For ease of comparison, the evaluation results of the six means of inquiry discussed in this chapter are given in Table 5.1 and illustrated in Figure 5.3. All of the means belong to one of the two highest relevance categories: The means in the first two sections are aimed at identifying evolutionary function/s of consciousness and those in the last two sections at discovering its evolutionary origin and development since. All of these evolutionary strategies have a high efficacy (7.0 to 8.5), though there are some concerns with the procedures of reverse engineering and the ontogenetic-parallels strategy. The practicality of the means of inquiry is moderate to high (6.0 to 7.5). Because none of them contain or are data collection methods, the means take published research studies as input, with attendant consequences on the available input and resource cost. The evolutionary strategies' promise scores span almost the entire range of the *high* label (6.7 to 8.4). Thus, despite some differences between them that are worth noting, they are all highly promising means of inquiry into the evolution of consciousness.

What, then, are the main factors that determine whether these highly promising evolutionary strategies reach their potential? First, the better a strategy is implemented, the better it tends to work. The quality of implementation in turn depends, among other things, on the expertise and competence of the researchers and the avail-

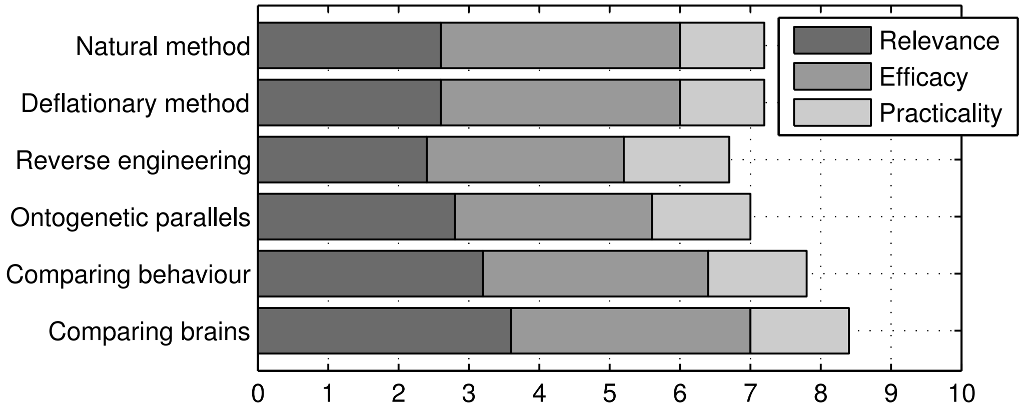

Figure 5.3. Dimensional promise profiles of the six evolutionary strategies evaluated in this chapter.

able resources. Second, the discussion of the comparative method has highlighted the centrality of reliable criteria for determining consciousness in other human and nonhuman animals. Third, the biggest limitation on the strategies' success is the inaccessibility of needed evidence, in particular historical evidence, as already pointed out in the introduction to this chapter. R. C. Richardson (2001, 2007) argued strongly against evolutionary psychology, narrowly construed, that evolutionary hypotheses about human mental traits will remain unconstrained speculation, because the required evidence will most likely never become available. However, even if an acceptable explanation of the evolution of consciousness by evolutionary biology's standards is unlikely, it is much too early to give up: The means of inquiry in the present chapter, along with the evidence that is or could soon be available, promise more than mere speculation. Finally, I have evaluated the strategies as they have been employed in consciousness research so far; employing more sophisticated versions of them should improve their success. I take a closer look at the evaluation and its implications in the next chapter.

6 Conclusion: Means of inquiry, the tool, and the future

The purpose of this thesis was to determine not only what means of inquiry are available for studying the biological evolution and function/s of consciousness but also how good they are or, more specifically, how promising they are with respect to the research goal of giving a scientifically respectable evolutionary explanation of consciousness. This required the construction of a suitable tool for evaluating the promise of means of inquiry. The evaluation tool I constructed has three dimensions—relevance, efficacy, and practicality—with two criteria each, which are assessed independently (except for the relevance criteria) and synthesised into dimensional and promise scores. This tool served to evaluate 23 means of inquiry that have been used in, or suggested to be used for, the investigation of the evolution of consciousness, including its adaptation status and evolutionary function/s.

The first task in this final chapter is to bring together the evaluation results from the three previous chapters in order to compare all evaluated means of inquiry and to derive recommendations for their use in consciousness research. To add to these recommendations, I then look to evolutionary psychology, as a wider field of the evolutionary study of mental traits, for methodological insights, warnings, and advice that can be usefully transferred to research into the evolution and function/s of consciousness. In the following section of this chapter, I examine the performance of the promise evaluation tool in the evaluations carried out in the three application chapters. The evaluation of the tool is important because it indicates how much confidence we can have in the evaluation results and, at least as importantly, how the evaluation tool can be improved for future use. The thesis ends with a summary of both the contributions it makes and my methodological recommendations for the evolutionary study of consciousness as well as for evaluations of means of inquiry.

6.1 Comparison of evaluated means of inquiry

When looking over the 23 evaluated means of inquiry, two things stand out even before consulting their evaluation results. The first is that they are all either arguments (Chapter 3) or strategies (Chapters 4 and 5). The second is that their aims, at least according to the evaluation, belong to one of the three evolutionary research subgoals. In particular, all arguments in Chapter 3 are aimed at showing that consciousness is an adaptation (an adaptation category), half of the strategies in Chapters 4 and 5 are aimed at identifying the evolutionary function/s of consciousness (what adaptation), and the other half at identifying its evolutionary development (origin and history). The an-adaptation arguments in Chapter 3 were grouped exactly according to these two criteria, an-adaptation and arguments, but the other means of inquiry were not selected in this way. I did, however, specify in section 2.4 that I would not evaluate means of inquiry aimed at showing that consciousness has a function, which explains their absence in the evaluation. I suspect that the popularity of evolutionary speculations has meant that even those means of inquiry whose proper domain is the identification of current function/s have been employed for evolutionary purposes. Assigning them to the highest reasonable relevance category is consistent with my instruction to do so, which was motivated by the present research goal. The fact that the remaining means are all more or less general strategies might be a sign of a relatively young field of inquiry but probably has more to do with trying to answer big historical questions.

The promise scores of the evaluated means of inquiry vary substantially, especially considering that these means have been deemed fit for purpose by some researchers: from the 4.2 of the argument from apparent adaptedness to the 8.4 of brain comparisons.[1] In addition, three arguments are completely unpromising: the argument from evolution, the generalisation argument, and the argument from evolutionary history, only the last one of which might fare better in future evaluations. As long as these arguments are unpromising, I advise against their use. But note that although researchers should prefer the most promising means of inquiry, they should also refrain from abandoning means with low promise scores prematurely. Low promise does not imply uselessness: Such means may, for instance, be appropriate in certain spe-

[1]For names and subevaluation scores of all evaluated means of inquiry, see Tables 3.1, 4.1, and 5.1 in the Conclusion sections of the respective chapters.

cific research situations or for triangulation. The large differences in the promise of means of inquiry (i.e., in their potential contribution under ideal circumstances and with present-day practicality) support my claim about the great differences in the usefulness of means of inquiry into the evolution and function/s of consciousness (see section 2.2). These perceived differences led me to conceive and pursue the evaluation project presented in this thesis.

Researchers might, quite likely, decide to focus their work on a particular subgoal and could then look for the most relevant means of inquiry in the corresponding category. The recommendations derived from the present evaluation for the three evolutionary subgoals are the following, with promise scores given in parentheses:

- ***An adaptation.*** The most promising arguments for showing that consciousness is an adaptation are the arguments from species comparisons (6.5), complex design (6.3), and biological cost (6.2).

- ***What adaptation.*** The most promising strategies for identifying evolutionary function/s of consciousness are contrastive analysis (7.2), exclusion (7.2), and the natural method / deflationary method (7.2).

- ***Origin and history.*** The most promising strategies for identifying the evolutionary origin of consciousness and its development since are brain comparisons (8.4), behaviour comparisons (7.8), and inferences from models or theories (7.3).

Whether researchers compare means of inquiry in relation to the research goal or a subgoal, they should consult the tables with subevaluation scores (see Footnote 1) and the accompanying text in order to choose one or more means of inquiry that are best suited to their research situation.

The discrepancy between the moderately to highly promising means of inquiry and their mostly unconvincing applications in the literature has repeatedly come up. It would have been easy to criticise the specific applications one by one, but I wanted to pursue a more productive approach instead. By weeding out unpromising means of inquiry and rating others and by pointing out their strengths and weaknesses, researchers who are aware of this information are likely to make better method choices. The evaluation has shown that the means of inquiry can work well if they are applied properly. The challenge therefore is to take the evaluation results as an encouragement to try to come closer to the evaluated ideal circumstances, not to become (even more) blasé. In the next section, I mention other possible ways of improving the execution of evolutionary studies in consciousness research.

6.2 Learning from evolutionary psychology

Additional pointers for improving evolutionary studies of consciousness can come from the consideration of evolutionary psychology, broadly construed. Developments in this field of inquiry as well as methodological publications can highlight important issues and provide new ideas of how to address (aspects of) the evolution of consciousness. The suggestion to consult evolutionary psychology is just an instance of looking outside one's immediate expert area or focus of study to other areas with similar problems. In the face of frequent poor knowledge of evolutionary theory among consciousness researchers, I give some brief examples of fairly general pointers and reminders here, simply to highlight another source of information and, one could say, of hope:

1. Several articles in recent years have outlined how psychological adaptations can be identified. For example, Andrews et al. (2002) discuss six standards of evidence and mention additional criteria and techniques for identifying adaptations.

2. Because developing hypotheses is an essential part of science, and a wealth of ideas is good for an emerging research area, a story-telling situation may not be problematic. Rossano (2003) suggested that "it may even be a sign of healthy vigor" (p. 45). On the other hand, an unconstrained "proliferation of adaptive hypotheses" (Griffiths, 1996, p. 516) is likely to be counterproductive. It is thus imperative that at least some of the hypotheses, after having been advanced, are developed further.

3. It is worth repeating here that it is vital to consider alternative explanations, including different evolutionary hypotheses on the function/s of consciousness, nonevolutionary explanations of consciousness, epiphenomenal explanations and, importantly, hypotheses about the less conscious competitors over which conscious organisms could have been selected.

4. Evolutionary approaches to mind and behaviour offer methods which have not yet been applied to the evolutionary study of consciousness such as, for example, analyses of cost-benefit trade-offs.

5. Additional methodological help can come from modern evolutionary biology. For instance, methods for determining the level of adaptation could assist with identifying how specific potential adaptive aspects of consciousness are likely to be.

6. Multimethod approaches are called for by many contemporary evolutionary psychologists (e.g., Durrant & Haig, 2001; Schmitt & Pilcher, 2004), and they are particularly relevant to methodologically difficult areas such as the evolutionary study of consciousness. Findings from different methods should be combined because each method has its strengths and weaknesses.

There are more lessons to be learned from evolutionary psychology. A focus on the possible adaptive problems that consciousness may have solved, and showing that consciousness can be subject to natural selection, for instance, could lead to important insights.

6.3 Evaluation of the evaluation tool

To see how accurate the evaluation tool is, we would need to know the promise of the means of inquiry already, but we do not. If some results have been surprising to readers (one supervisor commented on the promise score of the ontogenetic-parallels strategy "This high?"), their task is to try to locate the error in the tool (e.g., criterion missing, inadequate rubric, bad synthesis procedure) or reconsider their own assessment. Here I focus on difficulties that have come to my attention while applying the tool to the 23 means of inquiry. Before listing some of them according to dimensions, it is worth pointing out that the evaluation tool generally seems to agree with our expectations. For example, comparisons of brain and of behaviour were most promising, and none of the arguments for consciousness being an adaptation scored higher than the strategies aimed at identifying the evolutionary function/s or history of consciousness.

- ***Relevance.*** The relevance dimension does not take enough account of how well-supported expected content is likely to be; the rubric does not mention it at all. In general, the weighting of this dimension seemed rather too low than too high. The fertility bonus is much less likely to be assigned to the higher relevance categories, and the less relevant means in Chapter 3 that have received such a bonus may not have deserved it; they seem to have received it for their input without contributing anything themselves. The relevance dimension thus needs some fine-tuning.

- ***Efficacy.*** One of the main difficulties in applying the evaluation tool is caused by the specifics of how I tried to evaluate arguments and strategies with the same tool. The principle of charity in specifying arguments is responsible for them scoring full points on the procedure, whereas strategies struggle much more with this point. My recommendation is either to not evaluate these different types of means of inquiry with the same tool or to combine the input and procedure criteria into one. However, this point only means that the input and procedure criteria are more or less reversed for arguments and strategies, but their efficacy scores are probably hardly affected.

- ***Practicality.*** The two criteria, data availability and resource cost, are not independent (i.e., the more readily available input is, the lower the resource cost of obtaining the input). I accepted this because it reflects their relationship, but there could be a solution that works better with the synthesis procedure.

It was to be expected that the tool would require some changes. However, none of these difficulties are insurmountable. After the suggested changes to the tool have been made, the improved tool could be trialled with the problematic cases in the present evaluation. Overall though, as stated above, the tool's performance in this first proper round is reasonable. Among such things as providing absolute and relative subevaluation scores with verbal explanations, the tool has revealed many important weak spots of the evaluated means of inquiry and was able to allay some specific concerns with them.

The satisfactory performance of the evaluation tool indicates not only that its use is practicable and worthwhile but also that the tool is worthy of further development. In addition to the specific improvements suggested above, further evidence concerning the tool's validity and reliability should be collected. Of particular importance in this respect is its application by independent researchers who are knowledgeable in evolutionary and consciousness studies. Support for the validity and reliability of the evaluation tool would increase its psychometric credibility and hence its acceptability among psychologists. However, although the acceptance and use of the tool and its results are key to the success of this research, and validity and reliability in general are essential properties of good research tools, it is also important to keep development efforts in line with the characteristics and purpose of the tool: The present tool's application is not meant to create binding recommendations for all researchers (with the exception of recommendations against using completely irrelevant or inefficacious means of inquiry); this is because researchers differ in their research situations, approaches, personal strategies, and interpretations, and because plurality of employed means of inquiry is desirable. Instead, the tool's findings are meant to provide a first indication of the promise of means of inquiry to those researchers who would benefit from such methodological advice.

6.4 Contributions and recommendations

The last section in this chapter gives an overview of the contributions and recommendations of the thesis.

Major contributions. The two main components of the thesis, the construction of the evaluation tool for means of inquiry and its subsequent application to arguments and strategies used in the evolutionary study of consciousness, form the basis of its major contributions. The components each contribute two different kinds of knowledge, namely knowledge for instrumental use and knowledge for conceptual use (Cousins & Shulha, 2006; Rich, 1977), which are also called *knowledge for action* and *knowledge for understanding*. An overview of the contributions of the thesis can thus be structured as follows:

- ***Tool–conceptual.*** To construct the evaluation tool, I needed to determine what it means for a means of inquiry to be promising in the present context. The answer makes up a theory of the promise of means of inquiry, that is, an evaluative theory (Scriven, 2007b), which underpins the evaluation tool. Note that this was necessary because there was no suitable evaluation system or theory available. As such, the tool will, once published, be a significant contribution to the literature on the evaluation of means of inquiry. Because the tool is not specific to the content area (except for the tool's relevance rubric, which depends on the research goal) but to the content area's level of development, the evaluative theory would need to be adapted if applied at a different level of development either in the same area or a different one.

- ***Tool–instrumental.*** The tool may be used by any researcher who wants to evaluate means of inquiry. By supplying a tool, I hope to encourage researchers to evaluate means of inquiry in their fields. However, unless the evaluation is at roughly the same level of generality and development of the field, the criteria will most likely need to be modified.[2]

- ***Evaluation–conceptual.*** The evaluation of the 23 evaluations of arguments and strategies in this thesis provides new knowledge about them and thus increases our understanding of their strengths and weaknesses on all evaluated criteria and, where additional criticisms were discussed, beyond. This knowledge may then be used instrumentally to improve the means of inquiry and to apply them better.

- ***Evaluation–instrumental.*** The promise evaluation results themselves, along with the resulting recommendations about which means of inquiry to use and for what subgoal, are meant to—and hopefully will—help advance research on the biological evolution and function/s of consciousness. Whether they will do so depends, among other things, on their dissemination and uptake.

Together, these are the major contributions of this thesis.

Summary of recommendations. The most important recommendations, covering present findings and future research, that have come out of this research are summarised below:

- ***Use of evaluation results.*** To select, justify, or judge a means of inquiry into the evolution or function/s of consciousness, consult the promise evaluation results, including the subevaluations in numerical values and words. Far from being ultimate truths, these evaluations should be considered points for discussion and critical thought.

[2] In an interactive presentation to members of the Aotearoa New Zealand Evaluation Association (Upper South Island regional symposium, Hanmer Springs, July 2010), the tool was a very useful starting point for discussing the overlap with essential criteria of evaluation methods.

- ***Means of inquiry.*** Improve the evaluated means of inquiry where possible and needed most, as shown by the evaluation. Look for additional means of inquiry to be used in the evolutionary study of consciousness and evaluate them. Review the evaluation of these means every decade, unless it seems expedient to do so sooner.
- ***Evaluation tool.*** Improve the evaluation tool according to the suggestions made in the previous section. For example, it would be worth checking its interrater reliability, which could suggest further improvements. Future research could also attempt to test an adapted version of it on a past research situation in order to compare the tool's promise results with the actual outcome. (Note though that the most promising means of inquiry are not necessarily the ones that succeed, but they should do well on average.)

My final recommendation is more personal, though with potentially important consequences on consciousness research. It has to do with the subjective centrality of consciousness, which may or may not be an evolutionary adaptation (I mentioned this possibility in section 3.2). I argued for the prevalence of a centrality (and adaptation) bias in Chapter 3. Given that this bias seems to negatively influence research, it might be worthwhile, for example, to remind ourselves every now and then that we each consist of much more than consciousness. And maybe an occasional perspective change (see Fig. 6.1) could help us to better apply those means of inquiry into the biological evolution of consciousness.

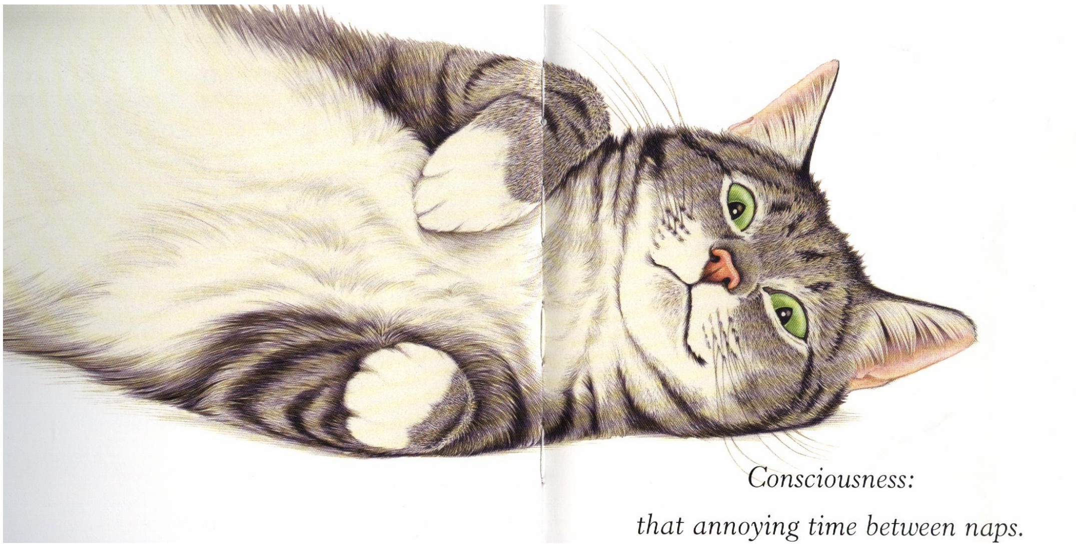

Figure 6.1. A different perspective on consciousness: that annoying time between naps. From *Purry Logic*, by J. Seabrook, 2008, Auckland, New Zealand: Hodder Moa. Copyright 2008 by Seabrook. Reprinted with permission.

References

Alcock, J. (2005). *Animal behavior: An evolutionary approach* (8th ed.). Sunderland, MA: Sinauer Associates.

Allen, C. (2010). Animal consciousness. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy* (Fall 2010 ed.). Retrieved from http://plato.stanford.edu/entries/consciousness-animal/

Allen, C., & Bekoff, M. (1997). *Species of mind: The philosophy and biology of cognitive ethology.* Cambridge, MA: MIT Press.

Amundson, R., & Lauder, G. V. (1994). Function without purpose: The uses of causal role function in evolutionary biology. *Biology and Philosophy*, *9*, 443–469.

Andrews, P. W., Gangestad, S. W., & Matthews, D. (2002). Adaptationism—How to carry out an exaptationist program. *Behavioral and Brain Sciences*, *25*, 489–504.

Århem, P., & Liljenström, H. (1997). On the coevolution of cognition and consciousness. *Journal of Theoretical Biology*, *187*, 601–612.

Århem, P., Lindahl, B. I. B., Manger, P. R., & Butler, A. B. (2008). On the origin of consciousness - some amniote scenarios. In H. Liljenström & P. Århem (Eds.), *Consciousness transitions: Phylogenetic, ontogenetic, and physiological aspects* (pp. 77–96). Amsterdam, The Netherlands: Elsevier.

Baars, B. J. (1986). *The cognitive revolution in psychology.* New York: Guilford Press.

Baars, B. J. (1988). *A cognitive theory of consciousness.* Cambridge, England: Cambridge University Press.

Baars, B. J. (1993). How does a serial, integrated and very limited stream of consciousness emerge from a nervous system that is mostly unconscious, distributed, parallel and of enormous capacity? In G. R. Bock & J. Marsh (Eds.), *Experimental and theoretical studies of consciousness* (pp. 282–290). Chichester, England: Wiley.

Baars, B. J. (1994). A thoroughly empirical approach to consciousness. *PSYCHE*, *1*. Retrieved from http://www.theassc.org/journal_psyche

Baars, B. J. (1997a). Contrastive phenomenology: A thoroughly empirical approach to consciousness. In N. Block, O. Flanagan, & G. Güzeldere (Eds.), *The nature of consciousness: Philosophical debates* (pp. 187–201). Cambridge, MA: MIT Press.

Baars, B. J. (1997b). *In the theater of consciousness: The workspace of the mind.* New York: Oxford University Press.

Baars, B. J. (1998). Metaphors of consciousness and attention in the brain. *Trends in Neurosciences*, *21*(2), 58–62.

Baars, B. J. (2002). The conscious access hypothesis: Origins and recent evidence. *Trends in Cognitive Sciences*, *6*(1), 47–52.

Baars, B. J. (2003a). How brain reveals mind: Neural studies support the fundamental role of conscious experience. *Journal of Consciousness Studies*, *10*(9–10), 100–114.

Baars, B. J. (2003b). Introduction: Treating consciousness as a variable: The fading taboo. In B. J. Baars, W. P. Banks, & J. B. Newman (Eds.), *Essential sources in the scientific study of consciousness* (pp. 1–10). Cambridge, MA: MIT Press.

Baars, B. J. (2005). Subjective experience is probably not limited to humans: The evidence from neurobiology and behavior. *Consciousness and Cognition*, *14*(1), 7–21.

Baars, B. J., & McGovern, K. (1996). Cognitive views of consciousness: What are the facts? How can we explain them? In M. Velmans (Ed.), *The science of consciousness: Psychological, neuropsychological and clinical reviews.* London: Routledge.

Badcock, C. R. (2000). *Evolutionary psychology: A critical introduction.* Cambridge, England: Polity Press.

Baldwin, J. M. (1896). Consciousness and evolution. *Psychological Review*, *3*(3), 300–309.

Baldwin, J. M. (1906). *Mental development in the child and the race: Methods and processes* (3rd ed.). Retrieved from http://www.brocku.ca/MeadProject/

Banks, W. P. (1993). Problems in the scientific pursuit of consciousness. *Consciousness and Cognition*, *2*(4), 255–263.

Banks, W. P., & Farber, I. (2003). Consciousness. In A. F. Healy & R. W. Proctor (Eds.), *Handbook of psychology: Vol. 4. Experimental psychology* (pp. 3–31). New York: Wiley.

Barlow, H. B. (1980). Nature's joke: A conjecture on the biological role of consciousness. In B. D. Josephson & V. S. Ramachandran (Eds.), *Consciousness and the physical world* (pp. 81–90). Oxford: Pergamon Press.

Barlow, H. B. (1987). The biological role of consciousness. In C. Blakemore & S. Greenfield (Eds.), *Mindwaves: Thoughts on intelligence, identity, and consciousness* (pp. 361–374). Oxford: Blackwell.

Bechtel, W., & Richardson, R. C. (1983). Consciousness and complexity: Evolutionary perspectives on the mind–body problem. *Australasian Journal of Philosophy*, *61*(4), 378–395.

Bering, J. M., & Bjorklund, D. F. (2007). The serpent's gift: Evolutionary psychology and consciousness. In P. D. Zelazo, M. Moscovitch, & E. Thompson (Eds.), *The Cambridge handbook of consciousness* (pp. 597–629). Cambridge, England: Cambridge University Press.

Blackmore, S. J. (2001). State of the art: Consciousness. *The Psychologist*, *14*(10), 522–525.

Blackmore, S. J. (2002). There is no stream of consciousness. *Journal of Consciousness Studies*, *9*(5–6), 17–28.

Blakemore, C. (2008, April). *The paradox of consciousness.* Paper presented at the satellite symposium The cognitive neuroscience of mind: A tribute to Michael Gazzaniga of the annual meeting of the Cognitive Neuroscience Society, San Francisco, CA.

Blaxter, L., Hughes, C., & Tight, M. (2006). *How to research* (3rd ed.). Maidenhead, England: Open University Press.

Block, N. (1993). [Review of the book *Consciousness explained*, by D. C. Dennett]. *The Journal of Philosophy*, *90*(4), 181–193.

Block, N. (1995). On a confusion about a function of consciousness. *Behavioral and Brain Sciences*, *18*, 227–247.

Brandon, R. N. (1990). *Adaptation and environment.* Princeton, NJ: Princeton University Press.

Breakwell, G. M., Hammond, S., Fife-Schaw, C., & Smith, J. A. (2006). Preface introducing research methods in psychology. In G. M. Breakwell, S. Hammond, C. Fife-Schaw, & J. A. Smith (Eds.), *Research methods in psychology* (3rd ed., pp. xviii–xxi). London, England: Sage.

Bridgeman, B. (2003). *Psychology and evolution: The origins of mind.* Thousand Oaks, CA: Sage.

Bringsjord, S. (1997). Consciousness by the lights of logic and commonsense. *Beha-*

vioral and Brain Sciences, *20*, 144–146. doi:10.1017/S0140525X97220054

Bringsjord, S., Noel, R., & Ferrucci, D. (2002). Why did evolution engineer consciousness? In J. H. Fetzer (Ed.), *Consciousness evolving* (pp. 111–138). Amsterdam: John Benjamins.

Broad, C. D. (1925). *The mind and its place in nature*. London, England: Routledge & Kegan Paul.

Buller, D. J. (1998). Etiological theories of function: A geographical survey. *Biology and Philosophy*, *13*(4), 505–527.

Buller, D. J. (2005). *Adapting minds: Evolutionary psychology and the persistent quest for human nature*. Cambridge, MA: MIT Press.

Buss, D. M. (1995). Evolutionary psychology: A new paradigm for psychological science. *Psychological Inquiry*, *6*(1), 1–30.

Buss, D. M. (2005). Introduction: The emergence of evolutionary psychology. In D. M. Buss (Ed.), *The handbook of evolutionary psychology* (pp. xxiii–xxv). Hoboken, NJ: Wiley.

Buss, D. M., Haselton, M. G., Shackelford, T. K., Bleske, A. L., & Wakefield, J. C. (1998). Adaptations, exaptations, and spandrels. *American Psychologist*, *53*(5), 533–548.

Butler, J. W. (2006). *Rethinking introspection: How we know (and fail to know) our own minds* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 3206972)

Byrne, R. W., Barnard, P. J., Davidson, I., Janik, V. M., McGrew, W. C., Miklósi, Á., & Wiessner, P. (2004). Understanding culture across species. *Trends in Cognitive Sciences*, *8*(8), 341–346.

Cabanac, M., Cabanac, A. J., & Parent, A. (2009). The emergence of consciousness in phylogeny. *Behavioural Brain Research*, *198*(2), 267–272.

Cairns-Smith, A. G. (1999). If qualia evolved . . . In S. R. Hameroff, A. W. Kaszniak, & D. J. Chalmers (Eds.), *Toward a Science of Consciousness III: The third Tucson discussions and debates* (pp. 271–279). Cambridge, MA: MIT Press.

Calvin, W. H. (1991). *The ascent of mind: Ice age climate and the evolution of intelligence*. New York: Bantam.

Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin*, *56*, 81–105. doi:10.1037/h0046016

Campbell, N. (2001). What was Huxley's epiphenomenalism? *Biology and Philoso-*

phy, *16*(3), 357–375.

Canfield, J. (1964). Teleological explanation in biology. *British Journal for the Philosophy of Science*, *14*(56), 285–295.

Caporael, L. R. (2001). Evolutionary psychology: Toward a unifying theory and a hybrid science. *Annual Review of Psychology*, *52*(1), 607–628.

Carroll, J. S., & Johnson, E. J. (1990). *Decision research: A field guide.* Newbury Park, CA: Sage Publications.

Carroll, S. P., Hendry, A. P., Reznick, D. N., & Fox, C. W. (2007). Evolution on ecological time-scales. *Functional Ecology*, *21*, 387–393. doi:10.1111/j.1365-2435.2007.01289.x

Carruthers, P. (2000). The evolution of consciousness. In P. Carruthers & A. Chamberlain (Eds.), *Evolution and the human mind: Modularity, language and metacognition* (pp. 254–275). Cambridge, England: Cambridge University Press.

Cartmill, M. (2000). Animal consciousness: Some philosophical, methodological, and evolutionary problems. *American Zoologist*, *40*(6), 835–846.

Chalmers, D. J. (1995). Facing up to the problem of consciousness. *Journal of Consciousness Studies*, *2*(3), 200–219.

Chalmers, D. J. (1996). *The conscious mind: In search of a fundamental theory.* New York: Oxford University Press.

Chalmers, D. J. (1997). Availability: The cognitive basis of experience. *Behavioral and Brain Sciences*, *20*, 148–149.

Chalmers, D. J. (2004). How can we construct a science of consciousness? In M. S. Gazzaniga (Ed.), *The cognitive neurosciences* (3rd ed.). Cambridge, MA: MIT Press.

Christensen, M. S., Ramsøy, T. Z., Lund, T. E., Madsen, K. H., & Rowe, J. B. (2006). An fMRI study of the neural correlates of graded visual perception. *NeuroImage*, *31*(4), 1711–1725.

Churchland, P. M. (1985). Reduction, qualia, and the direct introspection of brain states. *The Journal of Philosophy*, *82*(1), 8–28.

Churchland, P. S. (1983). Consciousness: The transmutation of a concept. *Pacific Philosophical Quarterly*, *64*, 80–95.

Churchland, P. S. (1986). *Neurophilosophy: Towards a unified science of the mind-brain.* Cambridge, MA: MIT Press.

Churchland, P. S. (1988). Reduction and the neurobiological basis of consciousness.

In A. J. Marcel & E. Bisiach (Eds.), *Consciousness in contemporary science* (pp. 273–304). Oxford, England: Clarendon Press.

Churchland, P. S. (1996). The hornswoggle problem. *Journal of Consciousness Studies*, *3*(5–6), 402–408.

Cleeremans, A. (2005). Computational correlates of consciousness. *Progress in Brain Research*, *150*, 81–98.

Cleeremans, A., & Jiménez, L. (2002). Implicit learning and consciousness: A graded, dynamic perspective. In R. M. French & A. Cleeremans (Eds.), *Implicit learning and consciousness* (pp. 1–40). Hove, England: Psychology Press.

Colman, A. M. (2009). *A dictionary of psychology* (3rd ed.) [Electronic version]. Oxford, England: Oxford University Press.

Cooke, N. J. (1994). Varieties of knowledge elicitation techniques. *International Journal of Human–Computer Studies*, *41*, 801–849. doi:10.1006/ijhc.1994.1083

Corballis, M. C. (2007). The evolution of consciousness. In P. D. Zelazo, M. Moscovitch, & E. Thompson (Eds.), *The Cambridge handbook of consciousness* (pp. 571–595). Cambridge, England: Cambridge University Press.

Coryn, C. L. S. (2007). *Evaluation of researchers and their research: Toward making the implicit explicit* (Doctoral dissertation, Western Michigan University, Kalamazoo). Retrieved from http://researcheval.net/documents/Evaluation_of_Researchers_and_Their_Research_Toward_Making_the_Implicit_Explicit.pdf

Cousins, J. B., & Shulha, L. M. (2006). A comparative analysis of evaluation utilization and its cognate fields of inquiry: Current issues and trends. In I. F. Shaw, J. C. Greene, & M. M. Mark (Eds.), *Handbook of evaluation: Policies, programs and practices* (pp. 266–291). London: Sage.

Cowey, A., & Stoerig, P. (1995, January 19). Blindsight in monkeys. *Nature*, *373*, 247–249. doi:10.1038/373247a0

Crick, F., & Koch, C. (1990). Towards a neurobiological theory of consciousness. *Seminars in the Neurosciences*, *2*, 263–275.

Crick, F., & Koch, C. (1995). Are we aware of neural activity in primary visual cortex? *Nature*, *375*(6527), 121–123.

Crick, F., & Koch, C. (1998). Consciousness and neuroscience. *Cerebral Cortex*, *8*, 97–107.

Crook, J. H. (1980). *The evolution of human consciousness.* Oxford, England: Clarendon Press.

Cross, N. (1994). *Engineering design methods: Strategies for product design* (2nd ed.). Chichester, England: Wiley.

Crothers, L. M., & Levinson, E. M. (2004). Assessment of bullying: A review of methods and instruments. *Journal of Counseling & Development*, *82*, 496–503.

Cummins, R. (1975). Functional analysis. *The Journal of Philosophy*, *72*(20), 741–765.

Damasio, A. R. (1998). Investigating the biology of consciousness. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, *353*(1377), 1879–1882.

Darwin, C. R. (1859). *On the origin of species by means of natural selection, or the preservation of favoured races in the struggle for life.* London, England: John Murray.

Darwin, C. R. (1958). *The autobiography of Charles Darwin 1809–1882* (N. Barlow, Ed.). Retrieved from http://darwin-online.org.uk/ (Original work published 1887)

Darwin, C. R., & Wallace, A. R. (1858). On the tendency of species to form varieties; and on the perpetuation of varieties and species by natural means of selection. *Journal of the Proceedings of the Linnean Society of London. Zoology*, *3*, 45–62.

Davidson, E. J. (2005). *Evaluation methodology basics: The nuts and bolts of sound evaluation.* Thousand Oaks, CA: Sage.

Dawkins, R. (1985). *The blind watchmaker.* New York: Norton.

de Gelder, B., Tamietto, M., van Boxtel, G., Goebel, R., Sahraie, A., van den Stock, J. B., & . . . Pegna, A. (2008). Intact navigation skills after bilateral loss of striate cortex. *Current Biology*, *18*, R1128–R1129. doi:10.1016/j.cub.2008.11.002

Dehaene, S., Changeux, J.-P., Naccache, L., Sackur, J., & Sergent, C. (2006). Conscious, preconscious, and subliminal processing: A testable taxonomy. *Trends in Cognitive Sciences*, *10*(5), 204–211.

Dehaene, S., & Naccache, L. (2001). Towards a cognitive neuroscience of consciousness: Basic evidence and a workspace framework. *Cognition*, *79*, 1–37.

Dehaene, S., Naccache, L., Cohen, L., Bihan, D. L., Mangin, J.-F., Poline, J.-B., & Rivière, D. (2001). Cerebral mechanisms of word masking and unconscious repetition priming. *Nature Neuroscience*, *4*, 752–758. doi:10.1038/89551

Delacour, J. (1995). An introduction to the biology of consciousness. *Neuropsychologia*, *33*, 1061–1074. doi:10.1016/0028-3932(95)00048-8

Dennett, D. C. (1986). Julian Jaynes's software archeology. *Canadian Psychology / Psychologie Canadienne*, *27*, 149–154. doi:10.1037/h0080051

Dennett, D. C. (1987). Consciousness. In R. L. Gregory (Ed.), *The Oxford companion to the mind* (pp. 160–164). Oxford, England: Oxford University Press.

Dennett, D. C. (1991). *Consciousness explained.* Boston: Little, Brown, and Co.

Dennett, D. C. (1995a). *Darwin's dangerous idea: Evolution and the meanings of life.* New York, NY: Simon & Schuster.

Dennett, D. C. (1995b). The path not taken. *Behavioral and Brain Sciences*, *18*, 252–253.

Dessalles, J.-L. (2001). *Qualia and spandrels: An engineering perspective* (Technical Report ENST 2001-D-012). Retrieved from http://www.enst.fr/~jld/papiers/pap.cogni/Dessalles_01082301.pdf

Donald, M. (2001). *A mind so rare: The evolution of human consciousness.* New York: Norton.

Dooremalen, H. (2003). *Evolution's shorthand: A presentational theory of the phenomenal mind* (Doctoral thesis, University of Nijmegen, The Netherlands). Retrieved from http://webdoc.ubn.kun.nl/mono/d/dooremalen_h/evolsh.pdf

Dretske, F. I. (1997). What good is consciousness? *Canadian Journal of Philosophy*, *27*, 1–15.

Dunbar, G. (2005). *Evaluating research methods in psychology: A case study approach.* Malden, MA: Blackwell Publishing.

Dupoux, E., de Gardelle, V. R., & Kouider, S. (2008). Subliminal speech perception and auditory streaming. *Cognition*, *109*(2), 267–273.

Durrant, R., & Haig, B. D. (2001). How to pursue the adaptationist program in psychology. *Philosophical Psychology*, *14*(4), 357–380.

Edelman, D. B. (2009). Animal consciousness. In W. P. Banks (Ed.), *Encyclopedia of consciousness* (Vol. 1). London: Academic Press.

Edelman, D. B., Baars, B. J., & Seth, A. K. (2005). Identifying hallmarks of consciousness in non-mammalian species. *Consciousness and Cognition*, *14*(1), 169–187.

Edelman, D. B., & Seth, A. K. (2009). Animal consciousness: A synthetic approach. *Trends in Neurosciences*, *32*, 476–484. doi:10.1016/j.tins.2009.05.008

Edelman, G. M. (1989). *The remembered present: A biological theory of consciousness.* New York: Basic Books.

Edelman, G. M. (2003). Naturalizing consciousness: A theoretical framework. *Proceedings of the National Academy of Sciences of the United States of America*, *100*(9), 5520–5524.

Edelman, G. M., & Tononi, G. (2000). *A universe of consciousness: How matter becomes imagination.* New York: Basic Books.

Ehrlich, P. R. (2000). *Human natures: Genes, cultures, and the human prospect.* Washington, DC: Island Press.

Elbert, J. W. (2000). *Are souls real?* Armherst, NY: Prometheus Books.

Ericsson, K. A., & Simon, H. A. (1980). Verbal reports as data. *Psychological Review*, *87*(3), 215ff.

Evans, J. St. B. T. (2005). *How to do research: A psychologist's guide.* Hove, England: Psychology Press.

Fernald, R. D. (2008). Evolution of vertebrate eyes. In A. I. Basbaum, A. Kaneko, G. M. Shepherd, G. Westheimer (Series Eds.), R. H. Masland, & T. D. Albright (Vol. Eds.), *The senses: A comprehensive reference: Vol. 1. Vision I* (pp. 9–24). San Diego, CA: Academic Press.

Fetzer, J. H. (Ed.). (2002). *Consciousness evolving.* Amsterdam, The Netherlands: John Benjamins.

Fink, A. (2005). *Conducting research literature reviews: From the internet to paper* (2nd ed.). Thousand Oaks, CA: Sage.

Flanagan, O. J. (1991). *The science of the mind* (2nd ed.). Cambridge, MA: MIT Press.

Flanagan, O. J. (1992). *Consciousness reconsidered.* Cambridge, MA: MIT Press.

Flanagan, O. J. (1995a). Consciousness and the natural method. *Neuropsychologia*, *33*(9), 1103–1115.

Flanagan, O. J. (1995b). Deconstructing dreams: The spandrels of sleep. *The Journal of Philosophy*, *92*, 5–27.

Flanagan, O. J. (2000). *Dreaming souls: Sleep, dreams, and the evolution of the conscious mind.* New York, NY: Oxford University Press. (Original work published 1999)

Flanagan, O. J., & Polger, T. W. (1995). Zombies and the function of consciousness. *Journal of Consciousness Studies*, *2*(4), 313–321.

Fodor, J. (2004, March 4). You can't argue with a novel [Review of the book *Radiant cool: A novel theory of consciousness*, by D. E. Lloyd]. *London Review of Books*, *26*(5), 30–31.

Foundation for Young Australians, & Sharp, C. A. (2000). *START do-it-yourself eval-*

uation manual: An integrated approach to project management and evaluation [Online summary]. Retrieved from http://www.personalresearchandevaluation.com/documents/performance_measurement/START Web Version.pdf

Franklin, S. (2005). Evolutionary pressures and a stable world for animals and robots: A commentary on Merker. *Consciousness and Cognition*, *14*, 115–118.

Freeman, S., & Herron, J. C. (2007). *Evolutionary analysis* (4th ed.). Upper Saddle River, NJ: Pearson Prentice Hall.

Freud, S. (1933). *New introductory lectures on psycho-analysis.* London: Hogarth Press.

Frith, C. D., & Rees, G. (2007). A brief history of the scientific approach to the study of consciousness. In M. Velmans & S. Schneider (Eds.), *The Blackwell companion to consciousness* (pp. 9–22). Oxford: Blackwell Publishing.

Futuyma, D. J. (2005). *Evolution.* Sunderland, MA: Sinauer Associates.

Gangestad, S. W., & Simpson, J. A. (2007a). Methodological issues: The means of Darwinian behavioral science [Editors' introduction]. In S. W. Gangestad & J. A. Simpson (Eds.), *The evolution of mind: Fundamental questions and controversies* (pp. 23–28). New York, NY: Guilford Press.

Gangestad, S. W., & Simpson, J. A. (2007b). Whither science of the evolution of mind? In S. W. Gangestad & J. A. Simpson (Eds.), *The evolution of mind: Fundamental questions and controversies* (pp. 397–437). New York, NY: Guilford Press.

Ganten, D., Gerhardt, V., & Nida-Rümelin, J. (2008). *Funktionen des Bewusstseins* [Functions of consciousness]. Berlin, Germany: Walter de Gruyter.

Gärdenfors, P. (2008). Evolutionary and developmental aspects of intersubjectivity. In H. Liljenström & P. Århem (Eds.), *Consciousness transitions: Phylogenetic, ontogenetic, and physiological aspects* (pp. 281–305). Amsterdam, The Netherlands: Elsevier.

Gaulin, S. J. C., & McBurney, D. H. (2001). *Psychology: An evolutionary approach.* Upper Saddle River, NJ: Prentice Hall.

Gaulin, S. J. C., & McBurney, D. H. (2004). *Evolutionary psychology* (2nd ed.). Upper Saddle River, NJ: Prentice Hall.

German Psychological Society. (2005, June). *Empfehlungen der Deutschen Gesellschaft für Psychologie e.V. (DGPs) zur Einrichtung von Bachelor- und Masterstudiengängen in Psychologie an den Universitäten (Revision)* [Recommendations of the German Psychological Society for the establishment of bachelor's and master's programmes in psychology at the universities (revision)]. Retrieved from http://www.dgps.de/_download/2005/BMEmpfehlungDGPs-rev.pdf

Gillett, G. R. (1988). Consciousness and brain function. *Philosophical Psychology*, *1*(3), 327–341.

Ginsburg, S., & Jablonka, E. (2007a). The transition to experiencing: I. Limited learning and limited experiencing. *Biological Theory*, *2*, 218–230. doi:10.1162/biot.2007.2.3.218

Ginsburg, S., & Jablonka, E. (2007b). The transition to experiencing: II. The evolution of associative learning based on feelings. *Biological Theory*, *2*, 231–243. doi:10.1162/biot.2007.2.3.231

Godfrey-Smith, P. (2001). Three kinds of adaptationism. In S. Orzack & E. Sober (Eds.), *Adaptationism and optimality* (pp. 335–357). Cambridge: Cambridge University Press.

Goetz, A. T., & Shackelford, T. K. (2006). Modern application of evolutionary theory to psychology: Key concepts and clarifications. *American Journal of Psychology*, *119*(4), 567–584.

Golden, M. P. (Ed.). (1976). *The research experience.* Itasca, IL: Peacock.

Goldman, A. I. (1979). What is justified belief? In G. S. Pappas (Ed.), *Justification and knowledge: New studies in epistemology* (pp. 1–23). Dordrecht, The Netherlands: D. Reidel.

Goldman, A. I. (1986). *Epistemology and cognition.* Cambridge, MA: Harvard University Press.

Goldman, A. I. (1988). Strong and weak justification. *Philosophical Perspectives*, *2*, 51–69.

Goldman, A. I. (1993). Consciousness, folk psychology, and cognitive science. *Consciousness and Cognition*, *2*(4), 364–382.

Gopnik, A. (2010, January 3). Mind reading [Review of the book *Reading in the brain: The science and evolution of a human invention*, by S. Dehaene]. *The New York Times*, p. BR15. Retrieved from http://www.nytimes.com

Gould, S. J. (1977). *Ontogeny and phylogeny.* Cambridge, MA: Belknap Press. Retrieved from http://www.sjgarchive.org/library/library.htm

Gould, S. J. (1985). Challenges to Neo-Darwinism and their meaning for a revised view of human consciousness. In S. M. McMurrin (Ed.), *The Tanner lectures on human values* (Vol. 6, pp. 53–73). Salt Lake City: University of Utah Press.

Gould, S. J. (1991). Exaptation: A crucial tool for an evolutionary psychology. *Journal of Social Issues*, *47*(3), 43–65.

Gould, S. J. (1997, 9 October). Evolutionary psychology: An exchange [Reply]. *New York Review of Books*, *44*(15), 56–58. Retrieved from http://www.nybooks.com/articles/1070

Gould, S. J., & Lewontin, R. C. (1979). The spandrels of San Marco and the Panglossian paradigm: A critique of the adaptationist programme. *Proceedings of the Royal Society of London. Series B, Biological Sciences*, *205*(1161), 581–598.

Gould, S. J., & Vrba, E. S. (1982). Exaptation—A missing term in the science of form. *Paleobiology*, *8*(1), 4–15.

Graham, G. (1998). *Philosophy of mind: An introduction* (2nd ed.). Oxford, England: Blackwell.

Graham, G., & Horgan, T. E. (2002). Sensations and grain processes. In J. H. Fetzer (Ed.), *Consciousness evolving* (pp. 63–86). Amsterdam: John Benjamins.

Grantham, T., & Nichols, S. (1999). Evolutionary psychology: Ultimate explanations and panglossian predictions. In V. G. Hardcastle (Ed.), *Where biology meets psychology: Philosophical essays* (pp. 47–66). Cambridge, MA: MIT Press.

Gray, J. A. (1971). The mind-brain identity theory as a scientific hypothesis. *The Philosophical Quarterly*, *21*(84), 247–254.

Graziano, A. M., & Raulin, M. L. (2000). *Research methods: A process of inquiry* (4th ed.). Boston: Allyn and Bacon.

Green, A. S., Rafaeli, E., Bolger, N., Shrout, P. E., & Reis, H. T. (2006). Paper or plastic? Data equivalence in paper and electronic diaries. *Psychological Methods*, *11*, 87–105. doi:10.1037/1082-989X.11.1.87

Gregory, R. L. (1996a). Peculiar qualia. *Perception*, *25*(7), 755–756.

Gregory, R. L. (1996b). What do qualia do? *Perception*, *25*(4), 377–379.

Gregory, R. L. (1998). Brainy mind. *BMJ (British Medical Journal)*, *317*(7174), 1693–1695.

Griffiths, P. E. (1993). Functional analysis and proper functions. *British Journal for the Philosophy of Science*, *44*(3), 409–422.

Griffiths, P. E. (1996). The historical turn in the study of adaptation. *British Journal for the Philosophy of Science*, *47*(4), 511–532.

Griffiths, P. E. (2001). From adaptive heuristic to phylogenetic perspective: Some lessons from the evolutionary psychology of emotion. In H. R. Holcomb III (Ed.), *Conceptual challenges in evolutionary psychology: Innovative research strategies* (pp. 309–325). Dordrecht, The Netherlands: Kluwer Academic.

Grossenbacher, P. G. (2001). Multisensory coordination and the evolution of consciousness. In P. G. Grossenbacher (Ed.), *Finding consciousness in the brain: A neurocognitive approach* (pp. 277–314). Dordrecht, The Netherlands: John Benjamins.

Guba, E. G., & Lincoln, Y. S. (1981). *Effective evaluation: Improving the usefulness of evaluation results through responsive and naturalistic approaches.* San Francisco, CA: Jossey-Bass.

Güzeldere, G. (1995a). Consciousness: What it is, how to study it, what to learn from its history. *Journal of Consciousness Studies*, *2*(1), 30–51.

Güzeldere, G. (1995b). Problems of consciousness: A perspective on contemporary issues, current debates. *Journal of Consciousness Studies*, *2*(2), 112–143.

Güzeldere, G. (1995c). Varieties of zombiehood. *Journal of Consciousness Studies*, *2*(4), 326–333.

Güzeldere, G. (1997). The many faces of consciousness: A field guide. In N. Block, O. Flanagan, & G. Güzeldere (Eds.), *The nature of consciousness: Philosophical debates* (pp. 1–67). Cambridge, MA: MIT Press.

Haggard, P. (2005). Conscious intention and motor cognition. *Trends in Cognitive Sciences*, *9*(6), 290–295.

Haggard, P. (2008). Human volition: Towards a neuroscience of will. *Nature Reviews Neuroscience*, *9*(12), 934–946.

Haggard, P., & Eimer, M. (1999). On the relation between brain potentials and the awareness of voluntary movements. *Experimental Brain Research*, *126*(1), 128–133.

Haig, B. D. (2005). An abductive theory of scientific method. *Psychological Methods*, *10*(4), 371–388.

Hallett, M. (2007). Volitional control of movement: The physiology of free will. *Clinical Neurophysiology*, *118*(6), 1179–1192.

Hameroff, S. R. (1999). Evolution and function of consciousness—Introduction. In S. R. Hameroff, A. W. Kaszniak, & D. J. Chalmers (Eds.), *Toward a science of consciousness III: The third Tucson discussions and debates* (pp. 245–246). Cambridge, MA: MIT Press.

Hampton, S. (2009). *Essential evolutionary psychology.* London, England: Sage.

Harvey, P. H., & Pagel, M. D. (1991). *The comparative method in evolutionary biology.* Oxford, England: Oxford University Press.

Hayes, N., & Stratton, P. (2003). *A student's dictionary of psychology* (4th ed.).

London, England: Arnold.

Haynes, J.-D., Roth, G., Schwegler, H., & Stadler, M. A. (1998). Die funktionale Rolle des bewußt Erlebten [The functional role of conscious experience]. *Gestalt Theory*, *20*, 186–213.

Hebb, D. O. (1994). The problem of consciousness and introspection. In H. Gutfreund & G. Toulouse (Eds.), *Biology and computation: A physicist's choice* (pp. 831–846). Singapore, Singapore: World Scientific. (Reprinted from *Brain mechanisms and consciousness*, pp. 402–417, by J. F. Delafresnaye, Ed., 1954, Oxford, England: Blackwell)

Hempel, C. G. (1959). The logic of functional analysis. In L. Gross (Ed.), *Symposium on sociological theory* (pp. 271–307). Evanston, IL: Row, Peterson.

Henrich, J., Boyd, R., & Richerson, P. J. (2008). Five misunderstandings about cultural evolution. *Human Nature*, *19*, 119–137. doi:10.1007/s12110-008-9037-1

Henrich, J., & McElreath, R. (2003). The evolution of cultural evolution. *Evolutionary Anthropology*, *12*, 123–135. doi:10.1002/evan.10110

Herrick, C. J. (1945). The natural history of experience. *Philosophy of Science*, *12*(2), 57–71.

Heyes, C. M. (2000). Evolutionary psychology in the round. In C. M. Heyes & L. Huber (Eds.), *The evolution of cognition* (pp. 3–22). Cambridge, MA: MIT Press.

Hobson, J. A. (2009). REM sleep and dreaming: Towards a theory of protoconsciousness. *Nature Reviews Neuroscience*, *10*(11), 803–813.

Holcomb III, H. R. (2001). Introduction to part III. In H. R. Holcomb III (Ed.), *Conceptual challenges in evolutionary psychology: Innovative research strategies* (pp. 293–308). Dordrecht, The Netherlands: Kluwer Academic.

Humphrey, N. K. (1982). Consciousness: A just-so story. *New Scientist*, *95*, 473–477.

Humphrey, N. K. (1987a). The inner eye of consciousness. In C. Blakemore & S. Greenfield (Eds.), *Mindwaves: Thoughts on intelligence, identity, and consciousness* (pp. 377–381). Oxford, England: Blackwell.

Humphrey, N. K. (1987b). The uses of consciousness. *James Arthur Memorial Lecture*, American Museum of Natural History, New York. Available from http://cogprints.org/809/

Humphrey, N. K. (1992). *A history of the mind: Evolution and the birth of consciousness.* New York: Simon & Schuster.

Humphrey, N. K. (1995). The thick moment. In J. Brockman (Ed.), *The third culture:*

Beyond the scientific revolution (pp. 198–208). New York: Simon & Schuster.

Humphrey, N. K. (2000). The privatization of sensation. In C. M. Heyes & L. Huber (Eds.), *The evolution of cognition* (pp. 241–252). Cambridge, MA: MIT Press.

Humphrey, N. K. (2002). *The mind made flesh: Essays from the frontiers of psychology and evolution.* Oxford: Oxford University Press.

Humphrey, N. K. (2006). *Seeing red: A study in consciousness.* Cambridge, MA: Belknap Press of Harvard University Press.

Humphrey, N. K. (2008). Getting the measure of consciousness. *Progress of Theoretical Physics Supplement*, *173*, 264–269.

Huxley, T. H. (1874). On the hypothesis that animals are automata, and its history. *Nature*, *10*(253), 362–366.

Huxley, T. H. (1882). On the hypothesis that animals are automata, and its history. In *Science and culture and other essays* (pp. 199–245). London: Macmillan. (Reprinted from *Fortnightly Review*, 1874, *16*, 555–580)

Hyman, B. (1998). *Fundamentals of engineering design.* Upper Saddle River, NJ: Prentice Hall.

Hyslop, A. (2010). Other minds. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy* (Fall 2010 ed.). Retrieved from http://plato.stanford.edu/entries/other-minds/

James, W. (1879). Are we automata? *Mind*, *4*(13), 1–22.

James, W. (1890). *The principles of psychology* (Vol. 2). Retrieved from http://www.archive.org/details/principlesofpsyc002jame

James, W. (1907). *The principles of psychology* (Vol. 1). London: Macmillan. (Original work published 1890)

Jaynes, J. (1986). Consciousness and the voices of the mind. *Canadian Psychology / Psychologie Canadienne*, *27*, 128–148. doi:10.1037/h0080053

Jaynes, J. (2000). *The origin of consciousness and the breakdown of the bicameral mind.* Boston, MA: Houghton Mifflin. (Original work published 1976)

Jerison, H. J. (1985). On the evolution of mind. In D. A. Oakley (Ed.), *Brain and mind* (pp. 1–31). London: Methuen.

Johnson-Laird, P. N. (1988). *The computer and the mind: An introduction to cognitive science.* Cambridge, MA: Harvard University Press.

Joint Task Force of *Academic Medicine* and the GEA–RIME Committee. (2001). Review criteria for research manuscripts. *Academic Medicine*, *76*, 897–978.

Keenan, J. P., Rubio, J., Racioppi, C., Johnson, A., & Barnacz, A. (2005). The right hemisphere and the dark side of consciousness. *Cortex*, *41*, 695–704.

Keller, I., & Heckhausen, H. (1990). Readiness potentials preceding spontaneous motor acts: Voluntary vs. involuntary control. *Electroencephalography and Clinical Neurophysiology*, *76*(4), 351–361.

Kemp, S., & Strongman, K. T. (1994). Consciousness—A folk view. *New Zealand Journal of Psychology*, *23*(2), 42–47.

Kingsolver, J. G., & Pfennig, D. W. (2007). Patterns and power of phenotypic selection in nature. *BioScience*, *57*, 561–572. doi:10.1641/B570706

Kinsbourne, M. (2006). From unilateral neglect to the brain basis of consciousness. *Cortex*, *42*(6), 869–874. doi:10.1016/S0010-9452(08)70430-1

Kirk, R. (2009). Zombies. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy* (Fall 2010 ed.). Retrieved from http://plato.stanford.edu/entries/zombies/

Klahr, D., & Simon, H. A. (1999). Studies of scientific discovery: Complementary approaches and convergent findings. *Psychological Bulletin*, *125*, 524–543. doi:10.1037/0033-2909.125.5.524

Koch, C. (2004). *The quest for consciousness: A neurobiological approach*. Englewood, CO: Roberts and Company.

Kuhn, T. S. (1977). *The essential tension: Selected studies in scientific tradition and change*. Chicago, IL: University of Chicago Press.

Kulka, R. A. (1981). Idiosyncrasy and circumstance: Choices and constraints in the research process. *American Behavioral Scientist*, *25*, 153–178. doi:10.1177/000276428102500204

Lagercrantz, H. (2008). The emergence of consciousness in the newborn. In H. Liljenström & P. Århem (Eds.), *Consciousness transitions: Phylogenetic, ontogenetic, and physiological aspects* (pp. 161–176). Amsterdam, The Netherlands: Elsevier.

Laland, K. N., & Brown, G. R. (2002). *Sense and nonsense: Evolutionary perspectives on human behaviour*. Oxford: Oxford University Press.

Laland, K. N., & Hoppitt, W. (2003). Do animals have culture? *Evolutionary Anthropology: Issues, News, and Reviews*, *12*(3), 150–159.

Laland, K. N., Kendal, J. R., & Brown, G. R. (2007). The niche construction perspective: Implications for evolution and human behaviour. *Journal of Evolutionary Psychology*, *5*(1), 51–66.

Laland, K. N., Odling-Smee, J., & Feldman, M. W. (2000). Niche construction, biological evolution, and cultural change. *Behavioral and Brain Sciences*, *23*, 131–146.

Lamme, V. A. F. (2006). Towards a true neural stance on consciousness. *TRENDS in Cognitive Sciences*, *10*(11), 494–501.

Lau, H. C. (2008). Are we studying consciousness yet? In M. Davies & L. Weiskrantz (Eds.), *Frontiers of consciousness: The Chichele lectures* (pp. 245–258). Oxford, England: Oxford University Press.

Lau, H. C. (2009). Volition and the function of consciousness. In M. S. Gazzaniga (Ed.), *The cognitive neurosciences* (4th ed., pp. 1191–1200). New York, NY: MIT Press.

Laudan, L. (1977). *Progress and its problems: Towards a theory of scientific growth.* London, England: Routledge & Kegan Paul.

Laudan, L. (1981). A problem-solving approach to scientific progress. In I. Hacking (Ed.), *Scientific revolutions* (pp. 144–155). Oxford, England: Oxford University Press.

Laudan, L. (1984). *Science and values: The aims of science and their role in scientific debate.* Berkeley, CA: University of California Press.

Laudan, L. (1987). Progress or rationality? The prospects for normative naturalism. *American Philosophical Quarterly*, *24*, 19–31.

Laureys, S. (2005). The neural correlate of (un)awareness: Lessons from the vegetative state. *Trends in Cognitive Sciences*, *9*, 556–559. doi:10.1016/j.tics.2005.10.010

Laureys, S., Owen, A. M., & Schiff, N. D. (2004). Brain function in coma, vegetative state, and related disorders. *The Lancet Neurology*, *3*, 537–546. doi:10.1016/S1474-4422(04)00852-X

Leakey, R. E. (1994). *The origin of humankind.* London, England: Weidenfeld & Nicolson.

Leary, M. R. (2004). *Introduction to behavioral research methods* (4th ed.). Boston, MA: Allyn and Bacon.

Leary, M. R., & Buttermore, N. R. (2003). The evolution of the human self: Tracing the natural history of self-awareness. *Journal for the Theory of Social Behaviour*, *33*, 365–404. doi:10.1046/j.1468-5914.2003.00223.x

Lester, F. K., Jr. (1996). Editorial. *Journal for Research in Mathematics Education*, *27*, 130–132.

Lewontin, R. C. (1978). Adaptation. *Scientific American*, *239*(3), 156–169.

Libet, B. (1985). Unconscious cerebral initiative and the role of conscious will in voluntary action. *Behavioral and Brain Sciences*, *8*, 529–566.

Libet, B., Gleason, C. A., Wright, E. W., & Pearl, D. K. (1983). Time of conscious intention to act in relation to onset of cerebral activity (readiness-potential): The unconscious initiation of a freely voluntary act. *Brain*, *106*(3), 623–642.

Lieberman, D. A. (1979). Behaviorism and the mind: A (limited) call for a return to introspection. *American Psychologist*, *34*, 319–333. doi:10.1037/0003-066X.34.4.319

Liljenström, H., & Århem, P. (2008). *Consciousness transitions: Phylogenetic, ontogenetic, and physiological aspects*. Amsterdam: Elsevier.

Lipsey, M. W., & Wilson, D. B. (2001). *Practical meta-analysis*. Thousand Oaks, CA: Sage.

Lissitz, R. W. (Ed.). (2009). *The concept of validity: Revisions, new directions and applications*. Charlotte, NC: Information Age Publishing.

MacKinnon, D. P., Lockwood, C. M., Hoffman, J. M., West, S. G., & Sheets, V. (2002). A comparison of methods to test mediation and other intervening variable effects. *Psychological Methods*, *7*, 83–104. doi:10.1037/1082-989X.7.1.83

Macphail, E. M. (1998). *The evolution of consciousness*. Oxford: Oxford University Press.

Macphail, E. M. (2000). The search for a mental Rubicon. In C. M. Heyes & L. Huber (Eds.), *The evolution of cognition* (pp. 253–271). Cambridge, MA: MIT Press.

Macphail, E. M. (2008). A bird's eye view of consciousness. In H. Liljenström & P. Århem (Eds.), *Consciousness transitions: Phylogenetic, ontogenetic, and physiological aspects* (pp. 97–121). Amsterdam, The Netherlands: Elsevier.

Mandler, G. (1975). Consciousness: Respectable, useful, and probably necessary. In R. L. Solso (Ed.), *Information processing and cognition: The loyola symposium* (pp. 229–254). Hillsdale, NJ: Erlbaum.

Mandler, G. (2002). *Consciousness recovered: Psychological functions and origins of conscious thought*. Amsterdam, NL: John Benjamins Publishing Company.

Mangan, B. B. (1995). The dead hand: Commentary on Baars on contrastive analysis. *PSYCHE*, *1*. Retrieved from http://www.theassc.org/journal_psyche

Marcel, A. J. (1986). Consciousness and processing: Choosing and testing a null hypothesis. *Behavioral and Brain Sciences*, *9*, 40–41.

Marcel, A. J. (1988). Phenomenal experience and functionalism. In A. J. Marcel & E. Bisiach (Eds.), *Consciousness in contemporary science* (pp. 121–158).

Oxford: Oxford University Press.

Marcel, A. J. (2003). Introspective report: Trust, self-knowledge and science. *Journal of Consciousness Studies*, *10*(9–10), 167–186.

Martindale, C. (1977–1978). Theories of the evolution of consciousness. *Journal of Altered States of Consciousness*, *3*(3), 261–278.

Mathison, S. (1988). Why triangulate? *Educational Researcher*, *17*(2), 13–17. doi:10.3102/0013189X017002013

Maynard Smith, J. (1978). Optimization theory in evolution. *Annual Review of Ecology and Systematics*, *9*, 31–56.

Mayr, E. (1961). Cause and effect in biology. *Science*, *134*(3489), 1501–1506.

Mayr, E. (1983). How to carry out the adaptationist program? *The American Naturalist*, *121*(3), 324–334.

Mayr, E. (1994). Recapitulation reinterpreted: The somatic program. *Quarterly Review of Biology*, *69*, 223–232.

Mayr, E. (1996). The modern evolutionary theory. *Journal of Mammalogy*, *77*(1), 1–7.

Mayr, E. (2002). *What evolution is.* London, England: Weidenfeld & Nicolson. (Original work published 2001)

McGinn, C. (1989). Can we solve the mind–body problem? *Mind*, *98*(391), 349–366.

McGinn, C. (1991). *The problem of consciousness: Essays towards a resolution.* Oxford, England: Blackwell.

McGinn, C. (1999). *The mysterious flame: Conscious minds in a material world.* New York, NY: Basic Books.

McGrath, J. E. (1981). Dilemmatics: The study of research choices and dilemmas. *American Behavioral Scientist*, *25*, 179–210. doi:10.1177/000276428102500205

McMullin, E. (1976). The fertility of theory and the unit for appraisal in science. In R. S. Cohen, P. K. Feyerabend, & M. W. Wartofsky (Eds.), *Essays in memory of Imre Lakatos* (pp. 395–432). Dordrecht, The Netherlands: D. Reidel.

Mead, G. H. (1934). *Mind, self & society from the standpoint of a social behaviorist* (C. W. Morris, Ed.). Chicago, IL: University of Chicago Press.

Merker, B. H. (2005). The liabilities of mobility: A selection pressure for the transition to consciousness in animal evolution. *Consciousness and Cognition*, *14*, 89–114.

Merton, R. K. (1949). Science and democratic social structure. In R. K. Merton, *Social theory and social structure: Toward the codification of theory and research* (pp.

307–316). Glencoe, IL: Free Press. (Reprinted from “Science and technology in a democratic order,” 1942, *Journal of Legal and Political Sociology*, *1*, 115–126)

Mesoudi, A. (2007). Biological and cultural evolution: Similar but different. *Biological Theory*, *2*, 119–123. doi:10.1162/biot.2007.2.2.119

Mesoudi, A., Whiten, A., & Laland, K. N. (2006). Towards a unified science of cultural evolution. *Behavioral and Brain Sciences*, *29*, 329–347. doi:10.1017/S0140525X06009083

Mesulam, M.-M. (1998). From sensation to cognition. *Brain*, *121*, 1013–1052. doi:10.1093/brain/121.6.1013

Metzinger, T. (2003). *Being no one: The self-model theory of subjectivity.* Cambridge, MA: MIT Press.

Mignot, E. (2008). Why we sleep: The temporal organization of recovery. *PLoS Biology*, *6*(4), e106.

Miller, G. (2005). What is the biological basis of consciousness? *Science*, *309*(5731), 79.

Miller, G. A. (1966). *Psychology: The science of mental life.* Harmondsworth: Penguin Books. (Original work published 1962)

Mitchell, M. L., & Jolley, J. M. (2007). *Research design explained* (6th ed.). Belmont, CA: Thomson Wadsworth.

Mithen, S. (1999). Handaxes and ice age carvings: Hard evidence for the evolution of consciousness. In S. R. Hameroff, A. W. Kaszniak, & D. J. Chalmers (Eds.), *Toward a science of consciousness III: The third Tucson discussions and debates* (pp. 281–296). Cambridge, MA: MIT Press.

Moore, K. (2006). Biology as technology: A social constructionist framework for an evolutionary psychology. *Review of General Psychology*, *10*(4), 285–301.

Morsella, E. (2003). The function of phenomenal states: Is there progress for the ”softer problem” of consciousness? *Psychological Reports*, *93*, 435–440.

Morsella, E. (2005). The function of phenomenal states: Supramodular interaction theory. *Psychological Review*, *112*(4), 1000–1021.

Moulin, C. (2006). Consciousness. In G. Davey (Ed.), *Encyclopaedic dictionary of psychology* (p. 135). London: Hodder Arnold.

Myerson, J., Miezin, F., & Allman, J. (1981). Binocular rivalry in macaque monkeys and humans: A comparative study in perception. *Behaviour Analysis Letters*, *1*, 149–159.

Nagel, T. (1974). What is it like to be a bat? *Philosophical Review*, *83*(4), 435–450.

National Academy of Sciences & Institute of Medicine. (2008). *Science, evolution, and creationism*. Washington, DC: National Academies Press. Retrieved from http://www.nap.edu/catalog/11876.html

Nichols, S., & Grantham, T. (2000). Adaptive complexity and phenomenal consciousness. *Philosophy of Science*, *67*(4), 648–670.

Nickles, T. (1986). Remarks on the use of history as evidence. *Synthese*, *69*(2), 253–266.

Nickles, T. (1987a). Methodology, heuristics, and rationality. In J. C. Pitt & M. Pera (Eds.), *Rational changes in science* (pp. 103–132). Dordrecht, The Netherlands: D. Reidel.

Nickles, T. (1987b). 'Twixt method and madness. In N. J. Nersessian (Ed.), *The process of science: Contemporary philosophical approaches to understanding scientific practice* (pp. 41–67). Dordrecht, The Netherlands: Martinus Nijhoff Publishers.

Nickles, T. (1989). Heuristic appraisal: A proposal. *Social Epistemology*, *3*, 175–188. doi:10.1080/02691728908578530

Nickles, T. (2006). Heuristic appraisal: Context of discovery or justification? In J. Schickore & F. Steinle (Eds.), *Revisiting discovery and justification: Historical and philosophical perspectives on the context distinction* (pp. 159–182). Dordrecht, NL: Springer.

Nicolau, M. C., Akaârir, M., Gamundí, A., González, J., & Rial, R. V. (2000). Why we sleep: The evolutionary pathway to the mammalian sleep. *Progress in Neurobiology*, *62*(4), 379–406.

Nielsen, M., & Day, R. H. (1999). William James and the evolution of consciousness. *Journal of Theoretical and Philosophical Psychology*, *19*, 90–113.

Nilsson, D.-E. (2005). Photoreceptor evolution: Ancient siblings serve different tasks. *Current Biology*, *15*(3), R94–R96.

Nisbett, R. E., & Wilson, T. D. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review*, *84*(3), 231–259.

Nolan, D. (1999). Is fertility virtuous in its own right? *The British Journal for the Philosophy of Science*, *50*(2), 265–282.

Oakley, D. A. (1985). Animal awareness, consciousness and self-image. In D. A. Oakley (Ed.), *Brain and mind* (pp. 132–151). London: Methuen.

O'Hear, A. (1997). *Beyond evolution: Human nature and the limits of evolutionary* *explanation.* Oxford: Clarendon Press.

O'Keefe, J. (1985). Is consciousness the gateway to the hippocampal cognitive map? a speculative essay on the neural basis of mind. In D. A. Oakley (Ed.), *Brain and mind* (pp. 59–98). London: Methuen.

Overgaard, M. (2003). On the theoretical and methodological foundations for a science of consciousness. *Journal of Anthropological Psychology*, *13*, 2–13. Retrieved from http://mit.psy.au.dk/ap

Overgaard, M. (2004). Confounding factors in contrastive analysis. *Synthese*, *141*(2), 217–231.

Overgaard, M., Rote, J., Mouridsen, K., & Ramsøy, T. Z. (2006). Is conscious perception gradual or dichotomous? A comparison of report methodologies during a visual task. *Consciousness and Cognition*, *15*(4), 700–708.

Owen, J. M. (with Rogers, P. J.). (1999). *Program evaluation: Forms and approaches* (2nd ed.). St Leonards, NSW, Australia: Allen & Unwin.

Pahl, G., Beitz, W., Feldhusen, J., & Grote, K.-H. (2007). *Engineering design: A systematic approach* (3rd ed.; K. Wallace & L. Blessing, Eds. & Trans.). London, England: Springer.

Panksepp, J. (1998). *Affective neuroscience: The foundations of human and animal emotions.* New York, NY: Oxford University Press.

Panksepp, J., & Panksepp, J. B. (2000). The seven sins of evolutionary psychology. *Evolution and Cognition*, *6*(2), 108–131.

Pauen, M., Staudacher, A., & Walter, S. (2006). Epiphenomenalism: Dead end or way out? [Editors' introduction]. *Journal of Consciousness Studies*, *13*(1–2), 7–19. Retrieved from http://www.imprint.co.uk/pdf/13_1-2_Introduction.pdf

Penrose, R. (1987). Minds, machines and mathematics. In C. Blakemore & S. Greenfield (Eds.), *Mindwaves: Thoughts on intelligence, identity, and consciousness* (pp. 259–276). Oxford, England: Blackwell.

Pinker, S. (1997). *How the mind works.* New York, NY: Norton.

Pinker, S. (2004, April). *Is consciousness an evolutionary adaptation?* [Tape recording]. Paper presented at Toward a Science of Consciousness, Tucson, AZ. Berkeley, CA: Conference Recording Service.

Pinker, S. (2005). Foreword. In D. M. Buss (Ed.), *The handbook of evolutionary psychology* (pp. xi–xvi). Hoboken, NJ: Wiley.

Pinker, S. (2006). Deep commonalities between life and mind. In A. Grafen & M. Ridley (Eds.), *Richard Dawkins: How a scientist changed the way we think:*

Reflections by scientists, writers, and philosophers (pp. 130–141). Oxford: Oxford University Press.

Pockett, S. (2004). Does consciousness cause behaviour? *Journal of Consciousness Studies*, *11*(2), 23–40.

Polger, T. W. (2000a). Zombies. In M. Nani & M. Marraffa (Eds.), *A field guide to the philosophy of mind.* Retrieved from http://host.uniroma3.it/progetti/kant/field/zombies.htm

Polger, T. W. (2000b). Zombies explained. In D. Ross, A. Brook, & D. L. Thompson (Eds.), *Dennett's philosophy: A comprehensive assessment* (pp. 259–286). Cambridge, MA: MIT Press.

Polger, T. W. (2002). Consciousness, function of. In L. Nadel (Ed.), *Encyclopedia of cognitive science* (pp. 750–753). Basingstoke, England: Palgrave Macmillan.

Polger, T. W. (2007). Rethinking the evolution of consciousness. In M. Velmans & S. Schneider (Eds.), *The Blackwell companion to consciousness* (pp. 72–86). Oxford, England: Blackwell Publishers.

Polger, T. W., & Flanagan, O. J. (1999). Natural answers to natural questions. In V. G. Hardcastle (Ed.), *Where biology meets psychology: Philosophical essays* (pp. 221–247). Cambridge, MA: MIT Press.

Polger, T. W., & Flanagan, O. J. (2002). Consciousness, adaptation and epiphenomenalism. In J. H. Fetzer (Ed.), *Consciousness evolving* (pp. 21–41). Amsterdam, The Netherlands: John Benjamins.

Popper, K. R. (1945). *The open society and its enemies* (Vol. 2). London: Routledge.

Popper, K. R. (1977). Some remarks on panpsychism and epiphenomenalism. *Dialectica*, *31*, 177–186. doi:10.1111/j.1746-8361.1977.tb01363.x

Prinz, J. (2004). The fractionation of introspection. *Journal of Consciousness Studies*, *11*(7–8), 40–57.

Proctor, R. W., & Capaldi, E. J. (2001). Empirical evaluation and justification of methodologies in psychological science. *Psychological Bulletin*, *127*(6), 759–772.

Pyrczak, F. (2008). *Evaluating research in academic journals: A practical guide to realistic evaluation* (4th ed.). Glendale, CA: Pyrczak.

Ramachandran, V. S. (1980). Introduction. In B. D. Josephson & V. S. Ramachandran (Eds.), *Consciousness and the physical world* (pp. 1–15). Oxford: Pergamon Press.

Ramsøy, T. Z., & Overgaard, M. (2004). Introspection and subliminal perception. *Phenomenology and the Cognitive Sciences*, *3*, 1–23.

Reber, A. S. (1992). The cognitive unconscious: An evolutionary perspective. *Consciousness and Cognition*, *1*(2), 93–133.

Rechtschaffen, A. (1998). Current perspectives on the function of sleep. *Perspectives in Biology and Medicine*, *41*(3), 359–390.

Reeve, H. K., & Sherman, P. W. (1993). Adaptation and the goals of evolutionary research. *Quarterly Review of Biology*, *68*(1), 1–32.

Reichenbach, H. (1938). *Experience and prediction: An analysis of the foundations and the structure of knowledge.* Chicago, IL: University of Chicago Press.

Rensink, R. A. (2009, June). Towards a taxonomy of visual attention. In T. Schlicht (Chair), *Attention and consciousness.* Symposium conducted at the annual meeting of the Association for the Scientific Study of Consciousness, Berlin, Germany.

Rescher, N. (1977). *Methodological pragmatism: A systems-theoretic approach to the theory of knowledge.* Oxford, England: Blackwell.

Revonsuo, A. (2000). The reinterpretation of dreams: An evolutionary hypothesis of the function of dreaming. *Behavioral and Brain Sciences*, *23*, 877–901.

Revonsuo, A. (2006). *Inner presence: Consciousness as a biological phenomenon.* Cambridge, MA: MIT Press.

Rich, R. F. (1977). Uses of social science information by federal bureaucrats: Knowledge for action versus knowledge for understanding. In C. H. Weiss (Ed.), *Using social research in public policy making* (pp. 199–211). Lexington, MA: Lexington Books.

Richardson, M. K., & Keuck, G. (2002). Haeckel's ABC of evolution and development. *Biological Reviews*, *77*, 495–528. doi:10.1017/s1464793102005948

Richardson, R. C. (2001). Evolution without history: Critical reflections on evolutionary psychology. In H. R. Holcomb III (Ed.), *Conceptual challenges in evolutionary psychology: Innovative research strategies* (pp. 327–373). Dordrecht, The Netherlands: Kluwer Academic.

Richardson, R. C. (2007). *Evolutionary psychology as maladapted psychology.* Cambridge, MA: MIT Press.

Ridley, M. (2004). *Evolution* (3rd ed.). Malden, MA: Blackwell Publishing.

Roback, A. A. (1952). *History of American psychology.* New York: Library Publishers.

Robb, D., & Heil, J. (2009). Mental causation. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy* (Fall 2010 ed.). Retrieved from http://plato.stanford

.edu/entries/mental-causation/

Robinson, W. S. (2009). Epiphenomenalism. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy* (Fall 2010 ed.). Retrieved from http://plato.stanford.edu/entries/epiphenomenalism/

Rose, H., & Rose, S. P. R. (Eds.). (2000). *Alas, poor Darwin: Arguments against evolutionary psychology.* New York, NY: Harmony Books.

Rosenthal, D. M. (1986). Two concepts of consciousness. *Philosophical Studies*, *49*, 329–359.

Rosenthal, D. M. (1993). State consciousness and transitive consciousness. *Consciousness and Cognition*, *2*(4), 355–363.

Rosenthal, D. M. (2008). Consciousness and its function. *Neuropsychologia*, *46*(3), 829–840.

Rosenthal, D. M. (2009). Concepts and definitions of consciousness. In W. P. Banks (Ed.), *Encyclopedia of consciousness.* London: Academic Press.

Rosnow, R. L., & Rosenthal, R. (2005). *Beginning behavioral research: A conceptual primer* (5th ed.). Upper Saddle River, NJ: Pearson Prentice Hall.

Rossano, M. J. (2003). *Evolutionary psychology: The science of human behavior and evolution.* Hoboken, NJ: Wiley.

Rossi, P. H., Lipsey, M. W., & Freeman, H. E. (2004). *Evaluation: A systematic approach* (7th ed.). Thousand Oaks, CA: Sage.

Roth, G. (1999). Entstehen und Funktion von Bewußtsein [Origin and function of consciousness]. *Deutsches Ärzteblatt*, *96*, A-1957–A-1961. Retrieved from http://www.aerzteblatt.de/

Roth, G. (2000a). Entstehen und Funktion von Bewusstsein: Schlusswort [Origin and function of consciousness: Closing words]. *Deutsches Ärzteblatt*, *97*, A-544– A-545. Retrieved from http://www.aerzteblatt.de/

Roth, G. (2000b). The evolution and ontogeny of consciousness. In T. Metzinger (Ed.), *Neural correlates of consciousness: Empirical and conceptual questions* (pp. 77–97). Cambridge, MA: MIT Press.

Roth, G. (2001). The evolution of consciousness. In G. Roth & M. F. Wullimann (Eds.), *Brain evolution and cognition* (pp. 555–582). New York: Wiley.

Ruse, M. (2000). *The evolution wars: A guide to the debates.* Santa Barbara, CA: ABC-CLIO.

Schmitt, D. P., & Pilcher, J. J. (2004). Evaluating evidence of psychological adaptation: How do we know one when we see one? *Psychological Science*, *15*(10), 643–649.

Schooler, J. W. (2002). Re-representing consciousness: Dissociations between experience and meta-consciousness. *Trends in Cognitive Sciences*, *6*(8), 339–344.

Schooler, J. W., & Fiore, S. M. (1997). Consciousness and the limits of language: You can't always say what you think or think what you say. In J. D. Cohen & J. W. Schooler (Eds.), *Scientific approaches to consciousness* (pp. 241–257). Mahwah, NJ: Erlbaum.

Schwitzgebel, E. (2008). The unreliability of naive introspection. *Philosophical Review*, *117*, 245–273. doi:10.1215/00318108-2007-037

Scriven, M. (1972). The exact role of value judgments in science. *Proceedings of the Biennial Meeting of the Philosophy of Science Association*, *1972*, 219–247.

Scriven, M. (1991). *Evaluation thesaurus* (4th ed.). Newbury Park, CA: Sage.

Scriven, M. (1994). The final synthesis. *Evaluation Practice*, *15*(3), 367–382.

Scriven, M. (2001). Evaluation: Future tense. *American Journal of Evaluation*, *22*, 301–307. doi:10.1177/109821400102200303

Scriven, M. (2007a). *Key evaluation checklist.* Retrieved from http://www.wmich.edu/evalctr/checklists/kec_feb07.pdf

Scriven, M. (2007b). *The logic and methodology of checklists.* Retrieved from http://www.wmich.edu/evalctr/checklists/papers/logic&methodology_dec07.pdf

Seal, D. W., Bogart, L. M., & Ehrhardt, A. A. (1998). Small group dynamics: The utility of focus group discussions as a research method. *Group Dynamics: Theory, Research, and Practice*, *2*, 253–266. doi:10.1037/1089-2699.2.4.253

Searle, J. R. (1992). *The rediscovery of the mind.* Cambridge: MIT Press.

Searle, J. R. (1993). The problem of consciousness. *Consciousness and Cognition*, *2*(4), 310–319.

Searle, J. R. (1998). How to study consciousness scientifically. *Brain Research Reviews*, *26*, 379–387.

Searle, J. R. (2001). Free will as a problem in neurobiology. *Philosophy*, *76*, 491–514.

Sergent, C., & Dehaene, S. (2004). Is consciousness a gradual phenomenon? *Psychological Science*, *15*(11), 720–728.

Seth, A. (2009). Explanatory correlates of consciousness: Theoretical and computational challenges. *Cognitive Computation*, *1*, 50–63. doi:10.1007/s12559-009-9007-x

Seth, A. K. (2009). Functions of consciousness. In W. P. Banks (Ed.), *Encyclopedia of consciousness* (Vol. 1, pp. 279–293). London: Academic Press.

Seth, A. K., Baars, B. J., & Edelman, D. B. (2005). Criteria for consciousness in humans and other mammals. *Consciousness and Cognition*, *14*, 119–139.

Seth, A. K., Dienes, Z., Cleeremans, A., Overgaard, M., & Pessoa, L. (2008). Measuring consciousness: Relating behavioural and neurophysiological approaches. *Trends in Cognitive Sciences*, *12*(8), 314–321.

Shallice, T. (1972). Dual functions of consciousness. *Psychological Review*, *79*(5), 383–393.

Sheets-Johnstone, M. (1998). Consciousness: A natural history. *Journal of Consciousness Studies*, *5*, 260–294. Retrieved from http://www.imprint.co.uk/jcs.html

Shulman, R. G., Hyder, F., & Rothman, D. L. (2009). Baseline brain energy supports the state of consciousness. *Proceedings of the National Academy of Sciences*, *106*, 11096–11101. doi:10.1073/pnas.0903941106

Siegel, J. M. (2005, 27 October). Clues to the functions of mammalian sleep. *Nature*, *437*(7063), 1264–1271.

Sirigu, A., Daprati, E., Ciancia, S., Giraux, P., Nighoghossian, N., Posada, A., & Haggard, P. (2004). Altered awareness of voluntary action after damage to the parietal cortex. *Nature Neuroscience*, *7*(1), 80–84.

Sjölander, S. (1995). Some cognitive breakthroughs in the evolution of cognition and consciousness and their impact on the biology of language. *Evolution and Cognition*, *1*, 3–11.

Sjölander, S. (1997). On the evolution of reality—Some biological prerequisites and evolutionary stages. *Journal of Theoretical Biology*, *187*, 595–600. doi:10.1006/jtbi.1996.0392

Sjölander, S. (1999). How animals handle reality—The adaptive aspect of representation. In A. Riegler, M. Peschl, & A. von Stein (Eds.), *Understanding representation in the cognitive sciences* (pp. 277–281). New York, NY: Kluwer Academic/Plenum.

Sleutels, J. J. M. (2008, April). *Recent changes in the structure of consciousness?* Paper presented at Toward a Science of Consciousness, Tucson, AZ.

Smith, E. A. (2000). Three styles in the evolutionary analysis of human behavior. In L. Cronk, N. A. Chagnon, & W. Irons (Eds.), *Adaptation and human behavior: An anthropological perspective* (pp. 27–46). New York, NY: Aldine de Gruyter.

Snyder, F. (1966). Toward an evolutionary theory of dreaming. *American Journal of*

Psychiatry, *123*(2), 121–136.

Sober, E. (2000). Evolution and the problem of other minds. *The Journal of Philosophy*, *97*(7), 365–386.

Spencer, H. (1872). *The principles of psychology* (2nd ed., Vol. 1). Retrieved from http://www.archive.org/details/principlespsych03unkngoog

Sperry, R. W. (1964). *Problems outstanding in the evolution of brain function.* James Arthur lecture on the evolution of the human brain. New York, NY: The American Museum of Natural History.

Squires, E. (1990). *Conscious mind in the physical world.* Bristol: Adam Hilger.

Stearns, S. C., & Hoekstra, R. F. (2005). *Evolution: An introduction* (2nd ed.). Oxford: Oxford University Press.

Sterelny, K., & Griffiths, P. E. (1999). *Sex and death: An introduction to philosophy of biology.* Chicago: University of Chicago Press.

Sternberg, R. J. (Ed.). (2006). *Reviewing scientific works in psychology.* Washington, DC: American Psychological Association.

Stoerig, P. (2010). Cueless blindsight. *Frontiers in Human Neuroscience*, *3*, 74. doi:10.3389/neuro.09/074.2009

Stout, G. F. (1931). *Mind & matter: The first of two volumes based on the Gifford Lectures delivered in the University of Edinburgh in 1919 and 1921.* Cambridge, England: Cambridge University Press.

Sufka, K. J. (2000). Chronic pain explained. *Brain and Mind*, *1*, 155–179.

Sufka, K. J., & Turner, D. (2005). An evolutionary account of chronic pain: Integrating the natural method in evolutionary psychology. *Philosophical Psychology*, *18*, 243–257. doi:10.1080/09515080500169488

Sutherland, N. S. (1989). *Macmillan dictionary of psychology.* London: Macmillan.

Symons, D. (1979). *The evolution of human sexuality.* New York: Oxford University Press.

Tannenbaum, A. S. (2001). The sense of consciousness. *Journal of Theoretical Biology*, *211*(4), 377–391.

Terrace, H. S., & Metcalfe, J. (2005). *The missing link in cognition: Origins of self-reflective consciousness.* Oxford, England: Oxford University Press.

Thagard, P. (1989). Explanatory coherence. *Behavioral and Brain Sciences*, *12*, 435–467. doi:10.1017/S0140525X00057046

Thagard, P. (1992). *Conceptual revolutions.* Princeton, NJ: Princeton University

Press.

Tinbergen, N. (1963). On aims and methods of ethology. *Zeitschrift für Tierpsychologie*, *20*, 410–433.

Tong, F., Nakayama, K., Vaughan, J. T., & Kanwisher, N. (1998). Binocular rivalry and visual awareness in human extrastriate cortex. *Neuron*, *21*(4), 753–759.

Tononi, G., & Edelman, G. M. (1998). Consciousness and complexity. *Science*, *282*(5395), 1846–1951.

Tooby, J., & Cosmides, L. (1992). The psychological foundations of culture. In J. H. Barkow, L. Cosmides, & J. Tooby (Eds.), *The adapted mind: Evolutionary psychology and the generation of culture* (pp. 19–136). New York: Oxford University Press.

Tooby, J., & Cosmides, L. (2005). Conceptual foundations of evolutionary psychology. In D. M. Buss (Ed.), *The handbook of evolutionary psychology* (pp. 5–67). Hoboken, NJ: Wiley.

Trevena, J. A., & Miller, J. (2002). Cortical movement preparation before and after a conscious decision to move. *Consciousness and Cognition*, *11*(2), 162–190.

Tye, M. (1996). The function of consciousness. *Noûs*, *30*(3), 287–305.

Vaneechoutte, M. (2000). Experience, awareness and consciousness: Suggestions for definitions as offered by an evolutionary approach. *Foundations of Science*, *5*, 429–456.

Van Gulick, R. (1989). What difference does consciousness make? *Philosophical Topics*, *17*, 211–230.

Van Gulick, R. (1997). Understanding the phenomenal mind: Are we all just armadillos? part ii: The absent qualia argument. In N. Block, O. Flanagan, & G. Güzeldere (Eds.), *The nature of consciousness: Philosophical debates* (pp. 435–442). Cambridge, MA: MIT Press.

Van Gulick, R. (2009). Consciousness. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy* (Fall 2010 ed.). Retrieved from http://plato.stanford.edu/entries/consciousness/

Varela, F. J. (1996). Neurophenomenology: A methodological remedy for the hard problem. *Journal of Consciousness Studies*, *3*, 330–349.

Velicer, W. F., & Fava, J. L. (1998). Effects of variable and subject sampling on factor pattern recovery. *Psychological Methods*, *3*, 231–251. doi:10.1037/1082-989X.3.2.231

Velmans, M. (1991). Is human information processing conscious? *Behavioral and*

Brain Sciences, *14*, 651–668.

Velmans, M. (1993). A reflexive science of consciousness. In G. R. Bock & J. Marsh (Eds.), *Experimental and theoretical studies of consciousness* (Vol. CIBA Foundation symposium 174, pp. 81–91). Chichester, England: Wiley.

Velmans, M. (1994). A thoroughly empirical first-person approach to consciousness: Commentary on Baars on contrastive analysis. *PSYCHE*, *1*. Retrieved from http://www.theassc.org/journal_psyche

Velmans, M. (2007). An epistemology for the study of consciousness. In M. Velmans & S. Schneider (Eds.), *The Blackwell companion to consciousness* (pp. 711–725). Oxford, England: Blackwell Publishing.

Voneida, T. J. (1998). Sperry's concept of mind as an emergent property of brain function and its implications for the future of humankind. *Neuropsychologia*, *36*, 1077–1082. doi:10.1016/S0028-3932(98)00061-X

Wallace, A. R. (1889). *Darwinism: An exposition of the theory of natural selection with some of its applications* (2nd ed.). London: MacMillan.

Walter, S. (2007). Epiphenomenalism. In J. Fieser & B. Dowden (Eds.), *Internet encyclopedia of philosophy.* Retrieved January 13, 2010, from http://www.iep.utm.edu/epipheno/

Wegner, D. M. (2002). *The illusion of conscious will.* Cambridge, MA: MIT Press.

Weiskrantz, L. (1988). Some contributions of neuropsychology of vision and memory to the problem of consciousness. In A. J. Marcel & E. Bisiach (Eds.), *Consciousness in contemporary science* (pp. 183–199). Oxford, England: Clarendon Press.

Welshon, R. (2008). *Parietal cortex expansion, working memory, and the emergence of consciousness.* Paper presented at Toward a Science of Consciousness, Tucson, AZ.

Welshon, R. (in press). *Philosophy, neuroscience and consciousness.* Acumen.

West-Eberhard, M. J. (1992). Adaptation: Current usages. In E. F. Keller & E. A. Lloyd (Eds.), *Keywords in evolutionary biology* (pp. 13–18). Cambridge, MA: Harvard University Press.

Whewell, W. (1840). *The philosophy of the inductive sciences, founded upon their history* (Vol. 2). London: J. W. Parker.

Whiten, A., Horner, V., & Marshall-Pescini, S. (2003). Cultural panthropology. *Evolutionary Anthropology*, *12*(2), 92–105.

Whitt, L. A. (1992). Indices of theory promise. *Philosophy of Science*, *59*, 612–634.

Wilcke, J. C. (2003). *Functions of phenomenal consciousness* (Unpublished honours thesis). University of Canterbury, Christchurch, New Zealand.

Wilkes, K. V. (1984). Is consciousness important? *British Journal for the Philosophy of Science*, *35*(3), 223–243.

Wilkes, K. V. (1988). ——, yìshì, duh, um, and consciousness. In A. J. Marcel & E. Bisiach (Eds.), *Consciousness in contemporary science* (pp. 16–41). Oxford, England: Clarendon Press.

Williams, G. C. (1966). *Adaptation and natural selection: A critique of some current evolutionary thought.* Princeton, NJ: Princeton University Press.

Wilson, D. B., & Lipsey, M. W. (2001). The role of method in treatment effectiveness research: Evidence from meta-analysis. *Psychological Methods*, *6*, 413–429. doi: 10.1037/1082-989X.6.4.413

Wimsatt, W. C. (1976). Reductionism, levels of organization, and the mind–body problem. In G. G. Globus, G. Maxwell, & I. Savodnik (Eds.), *Consciousness and the brain: A scientific and philosophical inquiry* (pp. 205–267). New York, NY: Plenum Press.

Workman, L., & Reader, W. (2004). *Evolutionary psychology: An introduction.* Cambridge, England: Cambridge University Press.

Workman, L., & Reader, W. (2008). *Evolutionary psychology: An introduction* (2nd ed.). Cambridge, England: Cambridge University Press.

Wright, W. (2007). Explanation and the hard problem. *Philosophical Studies*, *132*(2), 301–330.

Zelazo, P. D., Gao, H. H., & Todd, R. (2007). The development of consciousness. In P. D. Zelazo, M. Moscovitch, & E. Thompson (Eds.), *The Cambridge handbook of consciousness* (pp. 405–432). Cambridge, England: Cambridge University Press.

Zuk, M. (2002). A straw man on a dead horse: Studying adaptation then and now. *Behavioral and Brain Sciences*, *25*, 533–534.

Author Index