Chapter 3
Behavioral Investment Theory

> *To do anything—locate food, find a mate, reproduce, compose a sonata, solve an equation—you have to stay alive with enough surplus energy to perform the task at hand. Energy management drove the foundational adaptive design of all ancestral intelligence systems. All subsequent design features evolved as integrated augmentations of this core system—including the part that ultimately gives rise to your [conscious] mind.*
> *La Cerra & Bingham (2002, p. 4)*

So there you are on the couch watching TV when an Oreo® cookie ad activates in you a desire to get a glass of milk. It has been a long day, and you are feeling a little spent. A small calculation takes place—almost subconsciously—as you decide whether it is worth the effort to get up and pour yourself a glass. Finally, the thirst wins out. You pull yourself up and head over to the refrigerator. But scanning the contents you find no milk, resulting in a glance over at the trashcan, where you see the empty container. Feelings of irritation follow the interruption of your goal. The thought briefly enters your mind to head to the store, but it is quickly quashed—that would clearly require too much time and effort. You settle on a glass of orange juice, with mild feelings of annoyance.

Behavioral Investment Theory frames animal behaviors in terms of invested work effort, specifically expenditures of time and energy calculated in terms of costs and benefits. Watch any nature show documenting animal behavior and either implicitly or explicitly the narration will include references to animals making cost-benefit calculations about their activities. The other day I was watching a documentary on waterholes in Africa. Water can become a scarce and valuable resource for many animals, especially migrating mammals. But hungry crocodiles, lurking just beneath the surface, often patrol the waterholes. This particular episode vividly documented wildebeests tentatively approaching the water, ready to jump back at the slightest hint of danger. In the actions of the wildebeests, you could see the value of the water being weighed against the risk of a crocodile attack.

For an animal to survive and thrive, its actions must ultimately result in an overall positive return on its invested work effort. That is, it must ultimately acquire more workable energy and other necessary materials (i.e., particular chemicals) as a consequence of its behavior than the behavior costs. What mechanism allows animals

G. Henriques, *A New Unified Theory of Psychology*,
DOI 10.1007/978-1-4614-0058-5_3, © Springer Science+Business Media, LLC 2011

to calculate the costs of actions relative to the benefits? The short answer is the nervous system, which not only coordinates the behavior of the animal as a whole but also computes and predicts the values of certain behaviors relative to others. Framed as such, "the mind" (more on the meaning of this term later) is a decision-making system that calculates the value of the resources obtained and losses avoided, relative to the costs of spending the actions in the first place, the risks involved, and the value of other avenues of investment. The integrative potential of seeing the mind this way has been well articulated by Herb Gintis (2009), an economist who specializes in evolutionary biology and game theory. In some ways similar to the present work, he has attempted to develop a unified framework for the behavioral sciences, which for him includes economics, psychology, sociology, anthropology, and political science. He has argued that the central unifying principle underlying these disciplines is the view that the mind is a decision-making organ that calculates costs and benefits to arrive at choices. It is a formulation consistent with Behavioral Investment Theory.[1]

With its concept of investment explicit, Behavioral Investment Theory allows us to think about animal behaviors as a form of commerce with the environment. When we think of commerce, we think of spending something to receive something of value in return. In this case, it is the actions themselves that are spent—they cost the animal calories and time and, as was the case in the wildebeest example, often increase risk of injury or harm. Actions also can result in lost opportunities. If an animal is defending a territory, it is not finding food, and vice versa. Importantly, the famed radical behaviorist B. F. Skinner at one point characterized animal behavior as commerce (e.g., Skinner, 1938). Making the connection with behaviorism allows us to see one of the most important components of Behavioral Investment Theory, which is that the consequences the actions have shape the direction of future actions and allocations of mental resources. Think about this in monetary terms. If you purchase a shirt from a store only to have it come apart at the seams the first time you wear it, you are less likely to purchase that kind of shirt again. Likewise, if a fox spends time and energy hunting a skunk only to be sprayed with a noxious smelling substance, it will likely avoid such investments in the future.

The behavioral investment framework also makes clear that different environmental affordances will result in different levels of work effort. Major affordances or major dangers result in intense work effort to be acquired or avoided, smaller affordances or risks less so. If you have been exposed to major events—and almost

[1]It is worth noting that there are some important differences between Gintis' (2009) formulation and the one offered here. One major difference is the dimensions of complexity argument depicted by the ToK System. Another is the notion that the human mind consists of two separate systems of computation, a behavioral investment system and a justification system. Thus, while as an economist Gintis advocates for the rational actor model, as a psychologist, I advocate more for a rational emotional actor model, a point that will be made clearer as the book progresses. Nonetheless, it is important to note—as I have attempted to do throughout this book—connections between the unified theory and other integrative approaches.

all of us have been—you know how powerfully your investment system can become activated. When my son Jon started the first grade, he experienced some separation anxiety, and his despair at getting on the bus and going to school each morning made for a stressful couple of weeks during the beginning of the school year. One morning during this period I got up and noticed his door closed, which was surprising because he had been getting up early. When I went to wake him, and I found he was not in his bed. I called to my wife asking if he was downstairs, and she replied he was not. Upon hearing that, my whole system shifted into an enormously activated state. Images of him being kidnapped raced through my mind, as I went from room to room and he was nowhere to be found. Within a period of 20 seconds, I had gone from that sleepy feeling after just waking up to a state of intense panic, with blood coursing through my veins and my heart racing, as I ran through the house shouting for him. When he finally emerged tearfully from his hiding place underneath a blanket in his closet—after 2 minutes that felt like 20—relief poured through my body, and I had the strong desire to maintain proximity to him.

To get a basic idea of how we make investment calculations, think of the parallels in the examples mentioned thus far. In each case, input was referenced against some valued goal state, and this resulted in calculated work effort to achieve a desired outcome. In the milk example starting the chapter, the commercial made salient thirst drives that were then referenced relative to the work effort to get up off the couch. In the wildebeest example, the value of the water was referenced against the risk of attack, resulting in tentative and vigilant movements toward the waterhole. In the example concerning my son, the input that he was gone was referenced against my love for him and concern for his safety, and I was filled with energy and impulses to work toward anything that would return him to me.

With this basic understanding of Behavioral Investment Theory in mind, let me offer an outline for the remainder of this chapter. In the next section, I summarize the six foundational principles that together make up Behavioral Investment Theory. These six principles weave together key insights from the physical sciences, the biological sciences, and the various brain–behavior paradigms in an integrative manner that consolidates our knowledge. Although the principles are both commonly known and shared across several disciplines in animal behavioral science, Behavioral Investment Theory nevertheless is a novel proposal that can lead to new insights. After providing an overview of Behavioral Investment Theory, we then turn our attention to the concept of depression to show how it can help us define key constructs in human psychology.

There have been other proposals for unifying the mind and behavioral sciences that have much in common with Behavioral Investment Theory, and in the subsequent section I provide an overview of four such proposals. The authors of these proposals include a learning theorist who attempted to construct a unified theory of psychology, a cognitive psychologist with an evolutionary focus on the origin of the mind, a humanistically oriented computational engineer, and an evolutionary neuroscientist. By reviewing these proposals I hope to provide a fuller sense of Behavioral Investment Theory, and simultaneously show how it consolidates knowledge by providing a foundational frame of reference that allows one to see clearly how perspectives that are traditionally disparate can be effectively joined. In the final section of the chapter, I offer an integrative schematic depicting the architecture of the human mind as operating on four levels of information processing: (1) sensory-motor; (2) operant experiential; (3) imaginative thought; and (4) linguistic justification. The emphasis in the schematic is on how the information processing components of human psychological systems function as a whole.

The Six Principles of Behavioral Investment Theory

Behavioral Investment Theory consists of six foundational principles that are well established in the animal behavioral literature; however, they are often studied by different disciplines and paradigms that adopt different emphases and are often needlessly defined against one another or at the very least are disconnected from one another. As I review these principles, it is important to keep in mind that one of the primary advantages of Behavioral Investment Theory is that it consolidates knowledge and organizes disparate lines of thought into a coherent package that can be easily shared with broad application.

The first principle of Behavioral Investment Theory is the *principle of energy economics,* which is the notion that animals must, on the whole, acquire more workable energy from their behavioral investments than those behaviors cost. To explain this principle, I need to provide a little background on energy, entropy, and the laws of thermodynamics. Let's start with energy. Energy is the most fundamental substance in the universe, and it can be thought of as the ultimate common denominator. Matter, for example, is chunked frozen energy. Although energy can change forms, one of the most fundamental laws in physics—sometimes referred to as the first law of thermodynamics—is that energy is always conserved and cannot be created or destroyed. Thus according to modern physics, the amount of energy in the universe today is exactly the same as it was at the time of the Big Bang.

Physicists define energy in terms of the capacity to do work, and any complex system will have useful energy that can be directed to perform work, and useless energy that cannot be so directed. It is the complexity, order, and heat differentiation of a system that enables it to do work. In contrast, useless energy is random, disordered, and undifferentiated and cannot be directed to perform work. While the first law states that you cannot create or destroy energy, the second law of thermodynamics states that the availability of useful energy in a closed system (a system to which no new energy is added) will always decrease. This is why you can't build a perpetual motion machine; eventually, the workable energy must run down. Entropy is a measure of randomness and disorder, and it can also be thought of as the measure of useless or unworkable energy in a system. Thus another way of stating the second law of thermodynamics is that the entropy of a closed system will always increase.

To put these concepts in more concrete terms, think of a battery as a simple system that can do work. Energy went into constructing the battery, which resulted in the differentiation of charges. When a battery is being used, the flow of energy between the positive charge and negative charges produces potentially useful work. The battery also produces useless work in the form of heat, which is work that cannot be recaptured and redirected. Because of this essential fact, the battery will always become weaker—that is, its entropy will always increase—so long as no additional energy is added to the system.

What does the second law[2] have to do with animal behavior? Everything. Indeed, it has everything to do with biology in general. The reason is because living things exhibit an enormous amount of complex design and thus are very far from a state of thermodynamic equilibrium (or the state of maximum entropy). Thus, organic systems are intricately arranged in a manner that allows them to fend off the tendency toward randomness. At least until they die. Indeed, the decomposition that occurs following death is a testament to the pressures and powers of entropic processes. The complex functional design in organisms does not represent a violation of the second law because organisms are not closed systems. They import energy to maintain their ordered, differentiated, complex state. The process by which energy is acquired and distributed to fend off entropic decay can be considered one of the fundamental problems in biological design (e.g., Schrödinger, 1967).

La Cerra (2003) points out that what is true of biology is true of animal behavior as well. That is, as extensions of biological systems, neuro-behavioral systems must also be organized in a manner that allows them to solve the adaptive problem of entropy. La Cerra's model of the evolution of intelligence systems is based on this fundamental point and she describes the relationship well. She explains

> Plants are autotrophs, self-nourishing life forms. . .Because they are stationary, the resources plants require for photosynthesis must be available locally. Animals, on the other hand, are heterotrophs, life forms that depend on the consumption of plants, animals, or both in order to meet their bioenergetic requirements. Evolution has conferred upon animals the adaptive advantage of behavior—which enables them to forage for resources that are not locally available, to find shelter in order to conserve heat, and to otherwise manage energy in an effective manner. But there is a cost for this benefit: behavior itself requires energy. Consequently, animals have evolved intelligence systems that *function, first and foremost, as predictive bioenergetic cost/benefit analysis systems*. (pp. 442–443)

In other words, the capacity for movement was a great new advantage for animals but it also came at a cost in terms of energy expenditure. Because of the nature of energy economics, a fundamental principle of animal behavior is that the ratio of benefits acquired to costs incurred must, on the whole, be positive or else the animal will perish. La Cerra's (2003) article was titled *The First Law of Psychology Is the*

[2]The fundamental nature of the second law was well captured in a famous quotation by Sir Arthur Stanley Eddington, *The Nature of the Physical World* (1929):

> The law that entropy always increases holds, I think, the supreme position among the laws of Nature. If someone points out to you that your pet theory of the universe is in disagreement with Maxwell's equations—then so much the worse for Maxwell's equations. If it is found to be contradicted by observation—well, these experimentalists do bungle things sometimes. But if your theory is found to be against the second law of thermodynamics I can give you no hope; there is nothing for it but to collapse in deepest humiliation.

Second Law of Thermodynamics, which is a clever, alternative way of articulating the principle of energy economics.

Can we apply these insights in a manner that makes predictions? Yes. To offer just one example, a prediction from this analysis is that animals—given their knowledge and capacities—will spend the least amount of behavioral energy necessary to achieve the desired outcome. Think about this in terms of every day actions. When you are walking to class or heading to the store, you don't wander haphazardly in the general direction of where you want to go. Instead, you subconsciously calculate the least costly path to get to your destination. And when you look for them, you see that conservation of energy principles are omnipresent in animal behavior. For example, there is a powerful line of research in behavioral ecology called optimal foraging theory which shows how animals are exquisitely sensitive to assessing costs of foraging behavior relative to the risks of their behavior and the calories acquired from successful hunts.

To offer an example from that line of research, Richardson and Verbeek (1987) studied crows foraging on clams in the intertidal and noticed that they left quite a few clams behind after digging them up. A question arises from a Behavioral Investment Theory perspective. If they go to the trouble of digging them up in the first place why not eat them? Researchers noted that it was the small clams that the crows left behind and discovered that the answer to this question lies in handling time, which is the time it takes for the crows to open the clam. If the animal is assumed to be making behavioral investments so as to obtain a maximum return on energy and nutritional intake, a simple calculation can be performed that determines whether the gain of opening a small clam is optimal relative to the gain of opening a large clam plus the search and handling time. Richardson and Verbeek (1987) generated such a calculation and came up with a prediction for the percentage of clams eaten as a function of size that is remarkably similar to what the crows were observed to be doing. They were leaving small clams behind because the ratio of handling energy to energy consumed was less than that could be achieved by searching for a larger clam.

Hundreds of similar types of analyses have been conducted on animal foraging patterns, all consistent with the basic Behavioral Investment Theory framework. Indeed, the principle of energy economics has long been identified in various ways by a number of scientists. One of the clearest articulations was put forth by the economist George Zipf (1949) in *Human Behavior and the Principle of Least Effort*. The principle of least effort, defined by Zipf, means that a person

> will strive to solve his problems in such a way as to minimize the *total work* he must expend in solving both his immediate problems and his probable future problems. That in turn means that the person will strive to minimize *the probable average rate of his work expenditure* (over time). And in so doing he will be minimizing his effort. (italics in original) (p. 1)

Others have noted similar generalities in animal behavior. Gengerelli (1930) outlined the *principle of minima and maxima in animal learning*, which he stated was the fact that "The behavior of an organism elicited by a given situation which affords relief to an internal need tends, with repetition, to approach, in terms of [energy expended] the minimal limit compatible with the relief of that need." In *Principles of Behavior*, the Clark Hull developed the law of less work, which he articulated as follows: "If two or more behavior sequences, each involving a different amount of work, have been equally well reinforced an equal number of times, the animal will gradually learn to choose the less laborious behavior sequence leading to the attainment of the reinforcing state of affairs" (Hull, 1943, p. 294). These are all equivalent descriptions of the first general principle of Behavioral Investment Theory.

The second principle of Behavioral Investment Theory is the *evolutionary principle*. Whereas the first principle articulates the basic physical laws that must be operating for organisms in general and mental systems in particular to operate in and on the physical universe, the second principle is a statement of the processes by which these systems were built across the generations. The evolutionary principle of behavioral investment can be stated as follows: Genes that tended to build neuro-behavioral investment systems that expended behavioral energy in a manner that positively covaried with inclusive fitness were selected for, whereas genes that failed to do so were selected against. Thus, *inherited tendencies toward the behavioral expenditure of energy should be a function of ancestral inclusive fitness*.

This is a fairly standard evolutionary or neo-Darwinian formulation. Inclusive fitness is a term that is examined in greater detail in the chapter on the ToK System, when the unified theory is compared and contrasted with Edward O. Wilson's approach to unifying knowledge. Inclusive fitness is the term neo-Darwinian theorists use to refer to the totality of genetic material reproduced in the subsequent generations, and as will be explained later it means that we need to keep the reproductive success of kin in mind when we are thinking about selection pressures. The word *ancestral* is also important. It emphasizes the fact that evolution is neither a forward looking nor an intentional process; *we are not designed to spread our genes*. If we were, we would behave very differently. If my primary purpose in life was to leave as many genetic copies behind as possible, I would spend much more of my time at sperm banks and in other such adventures. Instead, our genetically endowed behavioral propensities are a function of the selection pressures of previous generations, what is commonly called the environment of evolutionary adaptation. That means that if the environment changes rapidly (i.e., within a few generations), there will likely be a mismatch between evolved behavioral predispositions and current environmental challenges. For example, it makes sense that we have strong preferences for rich, high calorie, fatty foods given the relative scarcity of such food during the Stone Age. But now that there is a McDonalds on almost every street corner in America, such preferences can easily result in maladaptive eating patterns (Allman, 1994).

It is now clear that the early grand theorists like Skinner, Freud, and Rogers did not effectively incorporate modern evolutionary theory into their frameworks. Although they were all strongly influenced by Darwin's theory, they did not systematically or accurately attend to the nature by which evolutionary forces would have fashioned the behavioral investment system. This is not terribly surprising because the modern evolutionary synthesis became well known after they had developed their primary insights. Nevertheless, it is a significant weakness in each of the grand theorists' models. Evolutionary psychologists like David Buss, Leda Cosmides, John Tooby, and Steven Pinker have made valuable contributions to human psychology by examining mental architecture in the light of modern evolutionary theory. We now know evolutionary processes have played a tremendous role in shaping the backbone of our neuro-behavioral investment system. That is, as a consequence of evolutionary forces, we are predisposed to respond to certain stimuli and prepared to learn certain associations. Recall the example of my reaction to my son Jon's temporary disappearance. The reaction was visceral, automatic, and reflected the basic architecture of my mammalian mental structure.

The third principle of Behavioral Investment Theory is the *principle of genetics*, which is the notion that genetic differences result in differences in behavioral investment systems. Research conducted in the field of behavioral genetics has demonstrated conclusively that a wide variety of different mental characteristics such as general intellectual abilities, personality traits like extraversion and conscientiousness, and susceptibility to various mental diseases like schizophrenia, bipolar disorder, and autism have a heritable component. Although the principle of genetics (which could also be called the principle of heredity) relates to the principle of evolution, it is important to be clear why they are quite separable. The principle of evolution is about the phylogeny of mental behavior, which refers to the various selection pressures that have operated across previous generations to shape the general architecture of the behavioral investment system. In contrast, the principle of genetics is about the genetic ontogeny of mental behavior, which is how the particular genetic combination formed at conception influences the development of the individuals' particular behavioral investment system. Thus, whereas the principle of evolution is intergenerational and applies to the species-level analyses, the principle of genetics is intragenerational (within the lifetime of the animal) and addresses individual differences. Figure 3.1 makes the relationship and difference between the

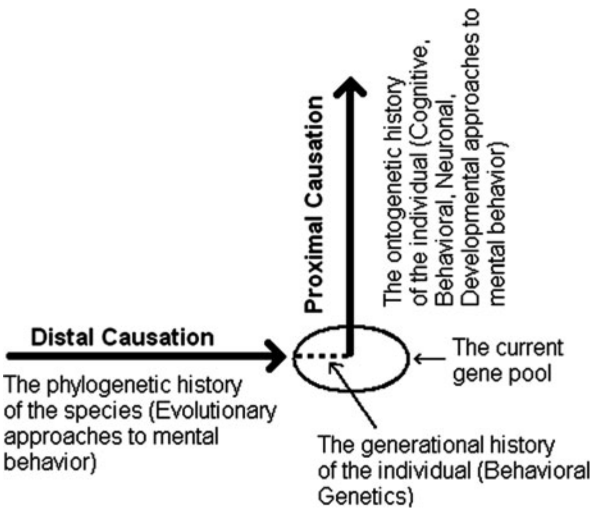

Fig. 3.1 The relationship between phylogeny and ontogeny

principle of evolution (distal causation) and the principle of genetics (which is a component of proximal causation) clear.

One additional point that needs to be made about the principle of genetics is that it is now clear that our genetic systems are not nearly as static as was once believed by most researchers. The historical conception has been of genetic programs unfolding to build organisms, with the causal arrow being pretty much one way. Research in developmental biology and other disciplines has now made clear that environmental events switch genes on and off (Oyama, 2000) and there are constant iterative forces between genes and the environment. Thus in many ways, the life of all organisms, even plants and organisms like bacteria, is a function of a transaction between the genetic system and the environment in a way that is far more fluid and bidirectional than is often supposed.

The fourth principle of Behavioral Investment Theory is the *computational control principle* and represents the central insight from cognitive (or computational) neuroscience, which is the idea that the nervous system is the organ of behavior and that it functions as an information processing system. Although now in common knowledge, it is important to recall that the notion that the nervous system is the organ of behavior represents a fairly recent discovery in the larger span of cultural history. The nervous system's role in animal behavior had been suspected at least since the time of the Greeks, and Descartes' (1596–1650) conception of animal and human behavior clearly involved the brain and nervous system. However, it wasn't until the middle of the nineteenth century that it became clear that the nervous system consisted of cells, and it wasn't until the middle of the twentieth century that consensus on how the nervous system functioned via networks of neural cells that could communicate via electrical impulses across synapses was achieved (Churchland, 1986).

This understanding of the structure of the nervous system combined with advances in information science to result in the computational theory of mind, which is the idea that the nervous system is an information processing system that works by translating physical and chemical changes in both the environment and in the body into neuronal patterns of information that represent the animal–environment relationship and compute action based on those representations. Consider, for example, when you see the color red what is actually happening is electromagnetic waves are entering through the lens of your eye and falling on the retina, which consists of various neural receptor cells (rods and cones) that respond to different physiochemical energy patterns. Your experience of red is a virtual world generated by patterns of neural firing. Red per se does not exist in the outside world, but instead is one of the ways your mind/brain system represents the world.

Although I chose the example of "red" so that you could see the relationship between physical information and mental representation, most of the information processed by the nervous system is not conscious. Cognitive scientists often differentiate conscious and nonconscious processes as explicit versus implicit. It is clear that most of the information processing that takes place is implicit; explicit processing is just the tip of the iceberg. Many studies utilizing a wide array of paradigms and methodologies clearly document the large number of implicit processes that are occurring as an individual experiences and considers her situation.

One compelling line of investigation clearly demonstrating the existence of implicit processing comes from neuroscience and is the example of blindsight (Weiskrantz, 1997). When the occipital lobe of the brain is severally damaged, an individual will lose all experience of sight. However, neuropsychologists discovered that even though patients report being completely blind, if asked to take a guess they nonetheless can accurately identify the location of an object placed on one side of the room or another. The reason is because certain implicit, visual processing tracts in the thalamus are still working. The control portion of the computational control principle is crucial, and in the section on the architecture of the human mind I will offer a schematic of the human mental system as a control system.

The fifth principle of Behavioral Investment Theory is the *learning principle*, and it pertains to how animal behavior evolves during the course of an animal's lifetime. In addition to the remarkable capacity for coordinated and purposeful movement, one of the most striking and well-documented aspects of animal behavior is the extent to which it is responsive to the changing demands of the environment over time. Animals learn to allocate their behavioral investments depending on the contingencies to which they are exposed. Processes of natural selection built the nervous system in a manner that allows it to be shaped by experiences. Although there are certainly some animal behaviors that are relatively fixed and hardwired, it is also the case that virtually all animals show some basic capacity to alter their responses to different contingencies, with some animals demonstrating remarkable flexibility in their behavioral repertoires. In describing how genes might build neuronal learning mechanisms, Richard Dawkins (1989) offered the following colorful description:

> One way for genes to solve the problem of making predictions in rather unpredictable environments is to build in the capacity for learning. Here the program may take the form of the following instructions to the survival machine: "Here is a list of things defined as rewarding: sweet taste in the mouth, orgasm, mild temperature, smiling child. And here is a list of nasty things: various sorts of pain, nausea, empty stomach, screaming child. If you should happen to do something that is followed by one of the nasty things, don't do it again, but on the other hand, repeat anything that is followed by the nice things". The advantage of this sort of programming is that it greatly cuts down the number of detailed rules that have to be built into the original program; and it is also capable of coping with changes in the environment that could not have been predicted in detail. (Dawkins, 1989, p. 57)

The learning principle can be stated succinctly as follows: behavioral investments that effectively move the animal toward animal–environment relationships that positively covaried with ancestral inclusive fitness are selected for (i.e., are reinforced), whereas behavioral investments that fail to do so are selected against and extinguished. A somewhat less formal way of stating the learning principle is the *pleasure pain parallel fitness principle.* In this formulation, pleasure (which includes all positive emotions) is nature's way of tagging benefits, whereas pain (which includes all negative emotions) is nature's way of tagging dangers and losses. The term fitness in this context has two meanings: one phylogenetic and the other ontogenetic. In the phylogenetic or evolutionary sense, it refers to the unconditioned or primary reinforcers that have been built into the system at the genetic level. In the ontogenetic or developmental sense, it refers to how pleasure and pain function to guide the animal to approach certain outcomes and avoid others and in the process acquire new behavioral repertoires. This relationship is explored in greater detail when discussing Arthur Staats' three function learning theory later in the chapter.

The foundational insight underlying the learning principle dates back to Thordike's Law of Effect; however, it was Skinner who saw clearly how crucial the process of behavioral selection was to understanding virtually all complicated animal behaviors. Although Skinner's radical behavioral philosophy was misguided because it falsely separated external from internal determinants (and tended to negate the latter), his notion of behavioral selection was a crowning intellectual insight and is foundational to all aspects of higher learning. Although most in psychology associate Skinner with concepts like reward and punishment, the idea of behavioral selection is really much more nuanced. The essence of Skinner's behavioral selection paradigm is that animals vary in the behaviors they emit, these varying behaviors have different consequences, and those consequences play a determinative role in the frequency, intensity, and duration of behaviors in the future. Most introductory psychology students learn about Skinner in terms of behavior modification. An educational example might be if a child is given a sticker for staying in their seat, they are more likely to stay in their seat in the future. While this is both true and helpful, I find the nature of behavioral selection to both be clearer and more compelling when I consider the manner in which I am constantly interacting with the environment and how closely and immediately my behaviors are tied to consequences.

For example, here I am typing away at my computer. When I hit a key, a consequence is a letter appears on the screen. My typing behavior is directly tied to this consequence. Consider what would happen, for example, if when I punched the key, no letter would pop up. First, I would likely push harder on the letters, and also vary which keys I hit. If those behaviors did not change the outcome, my typing behavior would quickly extinguish, and I would initiate another pattern of behavioral investment, like searching to see if the keyboard had come unplugged. The point here is that the letters on the computer screen are reinforcing my typing behavior, and if that reinforcer stops occurring, my typing behavior quickly alters. Fine-grained analyses of the relationship between actions and consequences are where you can see the magic of behavioral selection. Unfortunately, Skinner's paradigm is often taught in a very blunt way, with examples of how things like candy, money, or blame might shift someone's behavior. In actuality, we are constantly being shaped by the processes of behavioral selection. Moreover, many students don't realize that Skinner cared quite a lot about private thoughts and feelings; he just conceptualized them as forms of behavior. Either overtly or covertly we are constantly emitting a variety of mental behaviors, which in turn produce various internal or external consequences that either reinforce or extinguish future actions. In short, the behavioral selection paradigm helps us see clearly the way consequences shape the evolution of mental behavior.

The sixth and final principle of Behavioral Investment Theory is the *developmental principle*, which states that there are various genetically and hormonally regulated life history stages that require and result in different behavioral investment strategies. To provide a personal example, consider that when I was in the sixth grade most all of my friends were other boys, and I had little more than a passing interest in the opposite sex. By the time I was in the eighth grade, my investment value system had shifted dramatically, and much of my waking time was spent in daydreaming about possible encounters with girls I knew. What happened to make the shift? Was it that suddenly girls were interested in me, and I found such attention reinforcing? No, I can state with confidence that there had been little shift in the external patterns of reinforcement. Instead, what had happened was a cascade of hormonal releasers had dramatically shifted the basic structure of my investment value system; I had shifted into a new developmental stage called puberty.

LaCerra and Bingham (2002) describe in detail how animal behavioral investment systems—what they call intelligence systems—are structured in a developmentally sequenced way. They refer to such developmental structuring as the Life History Regulatory System (LHRS) and characterize it as follows:

> In addition to bottom line maintenance, life has an agenda of higher-order, sequential goals—major construction projects such as development, sexual maturation, and reproduction—that have to be scheduled. The agenda for this is in your genes—but it is flexible. These projects can't be launched without taking your individual life and your environmental circumstances into account. They're energy-expensive projects, and the LHRS is a strategic manager. It makes energetic trade-offs, allocating energy among competing goals. (La Cerra & Bingham, 2002, p. 25)

These six foundational principles of (1) energy economics, (2) evolution, (3) genetics, (4) computational control, (5) learning, and (6) development make up Behavioral Investment Theory, which in turn provides a unified, holistic framework for understanding animal behavior. It is important to note that this framework is very consistent with the work of Niko Tinbergen. Tinbergen was an ethologist who shared the 1973 Nobel Prize in Physiology/Medicine with Konrad Lorenz and Karl von Frisch for advances in the science of animal behavior. Tinbergen is well known for his work on fixed action patterns, which are fairly automatic ways animals respond to particular stimulus triggers. He is probably most famous for his framework of approaching animal behavior through "the four questions," which are as follows: (1) What are the mechanisms underlying animal behavior and how is the behavior elicited in relationship to recent learning? (2) What is the function of behavior in terms of survival and reproduction? (3) How did the behavior develop during the animal's life time? (4) How did the behavior evolve over the generations? I hope it is apparent to those familiar with Tinbergen's scheme that the principles of BIT correspond quite closely with the four questions. The principle of evolution corresponds directly with Tinbergen's fourth question. The principle of energy economics provides a conceptual framework to understand the ultimate function of behavior corresponding to the second question. The principles of learning and computational control correspond to question one. Finally, the principle of development corresponds to question three. The only principle that does not directly line up with Tinbergen's four questions is the principle of behavioral genetics, which essentially cuts across the four questions, in some ways binding them.

Given this overlap, it seems possible that some scholars who are familiar with animal behavioral research will consider the principles specified by Behavioral

Investment Theory as being well known. This criticism was indeed raised by Goertzen (2008, p. 841) who, although agreeing with the essence of Behavioral Investment Theory, added "I would go as far as to say that BIT is not exactly unique (or at the very least, is not overly surprising)." This raises the question of whether Behavioral Investment Theory is really necessary. Does it genuinely add anything to our current understanding? Yes, but let me first acknowledge that there is some truth to this criticism. For even if they are not explicit, many researchers in neuroscience, comparative psychology, ethology, behavioral ecology, and related disciplines already adopt similar foundational principles in their approach to animal behavior. Consequently, I imagine that the initial response of many animal behavioral researchers to Behavioral Investment Theory could well be, "Of course, there is nothing new here."

So what makes Behavioral Investment Theory valuable? In a nutshell, it provides a much needed consolidation across various domains of inquiry. This consolidation is crucial in two ways, one being epistemological and the other being in terms of intra- and inter-disciplinary communication. Epistemologically, Behavioral Investment Theory spells out a core of agreement and builds bridges between extant theoretical perspectives and consolidates existing knowledge. It achieves this consolidation better than previous approaches by virtue of the fact that it exists as part of the unified theory. In so doing, it plays a crucial role in a larger theoretical system that defines biology in relationship to psychology and psychology in relationship to the social sciences in a manner that allows the continuous and discontinuous aspects of the relationships between organisms (biology), animals (psychology), and humans (social sciences) to be made.

The second and related advantage of Behavioral Investment Theory is communication. As a clinical psychologist, I can confidently state that many professional psychologists are unaware of theoretical and empirical research programs in animal behavior. Sociobiology, ethology, behavioral ecology, and comparative psychology are distant disciplines that my colleagues in professional psychology tend to know little about. This is a shame because humans are, of course, animals, and any general principle about animals should also apply to humans. I believe Behavioral Investment Theory provides a simple way of communicating the key principles of animal behavior in a way that those who study and work with humans will be able to appreciate. To demonstrate the potential it has for linking an integrative frame for animal behavioral science with human problems, I will articulate how the logic of Behavioral Investment theory can help elucidate a better understanding of one of the most important constructs in clinical psychology and psychiatry, depression.

Applying the Insights of Behavioral Investment Theory: The Behavioral Shutdown Model of Depression

This section applies Behavioral Investment Theory to the construct of depression, which is an issue of tremendous public health importance. The World Health Organization currently ranks depression as the fourth largest contributor to the global burden of disease and estimates it will rank second by 2020. Epidemiological studies conducted in the United States have documented that depressive disorders are common. The estimate of lifetime prevalence of clinical depression is between 15 and 25% with higher rates for women and minorities (Gonzalez, Tarraf, Whitfield, & Vega, 2010). The 1-year prevalence of Major Depressive Disorder and dysthymia (a milder, but more chronic condition) among community residents is estimated to be approximately 10%, with an additional 11% of the population having significant subclinical symptoms (Zhang, Rost, & Fortney, 1999). These estimates suggest that in the United States alone, approximately 30 million people suffered from a depressive disorder and another 33 million suffered from substantial depressive symptoms in the past year. Research has demonstrated that depressive disorders are associated with difficulties in both biophysical and psychosocial functioning. MDD is associated with suicide and higher mortality rates in general, missed work, cognitive processing difficulties, and difficulties in social functioning. The direct and indirect costs associated with MDD in the US are estimated to be over $36 billion dollars annually, similar to costs associated with coronary heart disease (Hirschfeld et al., 1997). Thus, depression is a major public health issue.

And yet, despite its widespread importance and the enormous amount of research that has gone into the construct, there remains—as is the case with so many different concepts in psychology and psychiatry—significant debate regarding the precise nature of depression. As Ingram and Siegle note (2002, p. 87):

> The label "depression" has been used to discuss a mood state, a symptom, a syndrome. . . , a mood disorder, or a disease associated with biochemical or structural abnormalities. Although we may be tempted to ask, "Will the real depression please stand up?" the fact is that each of these constructs can legitimately lay claim to the term depression. Thus, depression is a construct that can mean very different things, which has important implications for decision making in research.

To get a flavor for why depression might mean different things to different researchers and how those different meanings might carry different sociopolitical implications, imagine two different television commercials. The first begins with an attractive woman isolating herself at a party. Everyone else appears to be having a good time, yet she stands in the background, ostensibly gripped in the throes of a seemingly inexplicable sadness. The cultural milieu is of upper middle class suburbia. A soft voice inquires and informs, "Have you experienced periods of depressed mood? Have you lost interest in things you used to enjoy? Do you feel tired, guilty, ineffective, or hopeless? Depression is an illness. Ask your doctor about new antidepressant treatments available." The implicit message of this commercial is clear. When people are suffering from depression, something has gone wrong with the physiology of the brain.

Now imagine a different commercial. This one begins with an impoverished woman getting slapped by her husband. Her three children are having difficulties in school. Her husband controls her, and she has little in the way of social support. She recently immigrated to the United States and cannot get a job because she only speaks a little English. She frequently faces prejudice and racism. The voice overlay asks, "Have you been feeling down or depressed, guilty or hopeless? Have you lost interest in things you usually enjoy? Depression is an illness. Ask your doctor about new antidepressant treatments available." Somehow the "depression as disease" message in this commercial is less convincing.

As these two vignettes illustrate, different portrayals can lead to radically different notions regarding the nature of depression. Yet how depression is conceptualized is critically important because the theoretical paradigms that guide our understanding of the condition influence public opinion, health policies, treatment strategies, and research. A Major Depressive Episode (MDE) is defined by the *Diagnostic and Statistical Manual-IV* (American Psychiatric Association, 1994) as the presence of five out of nine psychological and behavioral symptoms (depressed mood, anhedonia, agitation or retardation, fatigue or low energy, feelings of worthlessness or guilt, thoughts of death, change in appetite/weight, sleeping difficulties, and diminished ability to concentrate) every day for a period of 2 weeks. The prevailing model in psychiatry is that MDD is a disease of the brain (Judd, 1997), and there are, of course, good reasons for this position. In addition to the fact that depressive disorders are associated with difficulties in psychosocial functioning and higher mortality rates, neuroimaging studies have shown differences in the activity of the prefrontal cortex, the basal ganglia, the amygdala-hippocampus complex, and the thalamus in the brains of depressed individuals. Differences in the neuro-endocrine systems of depressed individuals have also been well documented. Additionally, Selective Serotonin Reuptake Inhibitors (SSRIs) are some of the most effective treatments for reducing depressive symptoms (see Krishnan & Nestler, 2010, for a recent review). Taken together, such findings form an impressive body of knowledge demonstrating that depressive disorders are associated with difficulties in functioning, that there are differences in the brain activity and brain chemistry of depressed individuals, and that psychopharmacological treatments are effective in reducing depressive symptoms.

And yet, despite these important findings, neurophysiological causal models of depression remain elusive. One possible reason for this failure is how depression is being conceptualized. When viewed as a disease, the psychological and behavioral symptoms that result in a diagnosis of MDD are generally assumed to be the product of neurophysiological dysfunctions. As such, differences in brain chemistry and/or brain activity between depressed individuals and controls are generally taken as evidence of brain pathology. However, it is important not to make a correlation–causation error when interpreting findings of associations between depression and psychosocial problems or physiological differences. A diagnosis of MDD is made on the presence of psychological and behavioral symptoms alone, not on etiology. The correct interpretation of the above findings is that depressive symptoms are correlated with difficulties in psychosocial functioning and differences in brain functioning. To infer that "depression" is the causal process underlying these difficulties requires one to make assumptions that are, at the very least, debatable. As is anecdotally illustrated by the second imaginary commercial, there are reasons to believe that depressive symptoms are often a reaction to difficulties in functioning. As will be argued below, it seems likely that depressive symptoms are both a cause and a consequence of difficulties in functioning in modern society.

Given this background, let's think about depression from the vantage point of Behavioral Investment Theory. The logic of Behavioral Investment Theory suggests that many instances of depression can be conceptualized as passive, avoidant behavioral strategies activated in response to situations that are chronically dangerous, humiliating, or repeatedly result in failure to achieve one's goals. Thus, instead of always being a result of biological malfunction, depression may instead reflect the basic structural design of the way the behavioral investment system operates, which is that animals will shut down when their behavioral investments consistently fail to effectively result in change. Let me explain.

Recall that, based on the principle of energy economics, Behavioral Investment Theory posits that mental behavior can be thought of as the process of expending energy or working in order to control and structure the animal–environment relationship, and that animals are disposed to expend energy in a way that corresponded to their ancestor's survival and reproductive success. Control of larger territories, access to better food, higher social status, etc. are obviously advantageous. However, the behavioral investment needed to acquire and maintain these resources can be expensive. It costs energy both in terms of basic calories and in terms of increasing risk of injury and loss. Resources are frequently not available or cannot be acquired, which means behavioral investments are fruitless. Additionally, competition over valuable resources can be fierce, often resulting in injury. This analysis corresponds to the energy economics principle of behavioral investment.

But what does this principle have to do with depression? Thinking of behavior being calculated on a cost-to-benefit ratio suggests that animals can maximize the ratio either by increasing benefits or by decreasing costs. Increasing benefits is associated with actively acquiring some resource (food, sex, status) in the environment via behavioral investment. The individual's state of actively working to increase benefit can be described as desire. Decreasing behavioral investment can also be a way in which animals deal with the cost-to-benefit ratio. There are many examples of behavioral shutdown mechanisms in nature, such as sleep, hibernation, and exhaustion, which function to decrease behavioral expenditure and conserve energy when resources are relatively scarce.

Broadly speaking, behavioral shutdown should result if an animal is consistently getting a poor return (i.e., high costs, little benefit) from its behavioral investment. That is, if an animal is spending eight behavioral units and only getting back four units of value, then that is a bad ratio. If it tries everything in its behavioral repertoire yet the ratio remains the same, a "best in a bad situation" solution is to decrease the amount of the behavioral investment in an effort to reduce net loss. It is better to expend two and get back one unit over the same period of time than the eight to four ratio previously obtained. This understanding gives rise to the Behavioral Shutdown Model which suggests that depression arises out of an evolved tendency to decrease behavioral expenditure in response to chronic danger, stress, or consistent failure to achieve one's goals (see also Beck, 1999a; Gilbert, 1998; Nesse, 2000). Put slightly different, according to the Behavioral Shutdown Model, we should think about depression as a state of behavioral shutdown.

The Behavioral Shutdown Model offers a potential explanation for many features of depression. For example, it strongly predicts that depression should be more likely to occur in situations that are chronically dangerous, humiliating, or repeatedly result in failure to achieve one's goals. These are circumstances in which the cost-to-benefit ratio is the worst and therefore the most effective strategy is to reduce costs. Consistent with this prediction, situations in which the individual feels chronically trapped or humiliated are most likely to produce symptoms of depression. To give just one example, almost 50% of battered women are depressed (Golding, 1999). There is also strong evidence that the onset of many Major Depressive Episodes are preceded by major stressful life events. Also consistent with the Behavioral Shutdown Model, rates of MDD vary with socioeconomic status. Those individuals in the lowest quartile of socioeconomic status are almost twice as likely to be depressed compared with those in the highest quartile (Yu & Williams, 1999).

In addition to offering an explanation as to why certain situations are more likely to result in depression, the Behavioral Shutdown Model also explains many of the symptoms of depression. The model explains why emotional pain is such a prominent feature of depression, as the pain is a signal that things are not going well. Additionally, behavioral shutdown is the antithesis of active behavioral investment, and thus the Behavioral Shutdown Model explains why anhedonia is such a fundamental characteristic of depressive conditions. It also directly accounts for why low energy is such a prominent complaint. The model also explains why negative cognitions are so prominent in depression. Cognitive theorists have clearly documented how depressed individuals are hypersensitive to any indications of loss, failure, or rejection. In direct accordance with the Behavioral Shutdown Model, recent cognitive models have conceptualized depressed individuals as investors with few resources who take risk-aversive strategies to avoid loss (Leahy, 1997). In short, the Behavioral Shutdown Model offers a potential explanation for many of the symptoms of depression.

The Behavioral Shutdown Model also provides explanations for findings that are difficult to explain from a disease model perspective. Because so many different things can result in difficulties in solving important problems, the model accounts for why so many different causal pathways result in depression. Behavioral shutdown should be a matter of degree, thus the Behavioral Shutdown Model also accounts for why symptoms of depression exist on a continuum that range from chronic, severe depressions to minor depressions to adjustment disorders to low mood. Since the model suggests depression should be associated with difficulties in functioning, the model explains why depressive symptoms evidence such a high comorbidity with other mental disorders, especially anxiety. Finally, because it is an evolutionary model, it also readily accounts for the fact that there is a substantial genetic component associated with depression.

The Behavioral Shutdown Model is valuable in that it links the causes and triggers with the effects and symptoms of depression in a logical sequence. To give just a few examples, the model predicts that because depressed individuals are focused on avoiding further loss, they should perceive more negative and pessimistic outcomes than those who are not depressed. Depressed individuals should also be risk aversive and tend to avoid potentially threatening stimuli. Likewise, depressed individuals should be hypersensitive to loss, failure, rejection, or physical pain. Because depressed individuals should be inclined to give up when faced with difficulty, such individuals should demonstrate a very low tolerance for frustration. Also, depressed individuals should exhibit diminished curiosity and explorative tendencies and should shun uncertainty, novelty, and sensation seeking. They should be very averse to conflict, particularly with others who are of equal or higher status. They should also engage in less social exchange. Depressed individuals should also demonstrate a decrease in behavioral activity. In short, the Behavioral Shutdown Model makes many clear, easily testable predictions about both the triggers and symptoms associated with depressive condition. If these predictions were not borne out by empirical data, then the model would be wrong.

In addition to offering a theoretical model that makes predictions and thus may lead to new empirical insights, the Behavioral Shutdown Model also offers a way to understand the various conceptual confusions that exist about the nature of depression highlighted in quotation by Ingram and Siegle (2002) offered at the beginning of this section. The essence of the point made by Ingram and Siegle is the question of whether the construct of depression should be thought of as a normal mood state, a psychological disorder, or a biological disease. The Behavioral Shutdown Model suggests that depression, including Major Depressive Episodes, should be considered a state of behavioral shutdown. *That is what depression is*. With this conception, then the question of whether depression is a normal mood, a psychological disorder, or a biological disease is found in the cause of the behavioral shutdown.

When an individual is depressed as an obvious consequence of serious loss or chronic frustration, we can see depression as a normal reaction. For example, if my wife and children were killed in a car accident, I may well experience a state of profound "shutdown" as my entire psychological system would need to become recalibrated given that as a consequence of the trauma my primary pathways of investment would be gone. Interestingly, the founders of the DSM seem to have recognized this when they offer bereavement as an exception to diagnosing depression if the loss occurred in the past 2 months. Yet, this exception for grief raises the question about the woman in the second imaginary commercial? Her shutting down also could be conceptualized as a very understandable reaction to the inability to find pathways of productive investment. However, as it currently stands, only bereavement exempts one from a diagnosis of MDD, whereas racism, poverty, isolation, and abuse do not.

In contrast, when we see depression arising as a function of vicious cycles of behavioral investment, where initial stressors lead to an ineffectual shutting down resulting to greater loss and this in turn leads to more depressive feelings, we can consider the condition a psychological disorder. In fact, this is essentially the behavioral conception of depression and consistent with a common behavioral treatment, called behavioral activation (Jacobson, Martell, & Dimidjian, 2001). The cognitive perspective similarly emphasizes how negative interpretations can lead to vicious depressive cycles. Or, from a more psychodynamic perspective, consider how self-criticisms so prominent in depressed individuals might sometimes function to justify submission and the inhibition of aggressive impulses, and this leads to greater shutdown. All of these models are conceptions of depression as a psychological disorder, whereby the shutdown is resulting in vicious cycles.

The perspective here also allows for a biopsychiatric conceptualization and clarifies the distinction between a disease and a behavioral disorder. As I have argued elsewhere (Henriques, 2002) the concept of "disease" is an applied biological construct, defined as a harmful breakdown in the function of an evolved mechanism. This construct can be conceptually differentiated from psychological disorders in which rigid, maladaptive behavioral patterns result from vicious behavioral cycles, as described above in the context of depression. Thus, severe depressive responses that occur in the absence of behavioral ineffectiveness or loss can be considered depressive diseases because such occurrences reflect a breakdown in the functioning of the basic bio-psychological architecture. In sum, Behavioral Investment Theory leads to a new conceptualization of depression as a state of behavioral shutdown which offers much potential to clarify how and why depression can legitimately be considered a normal reaction, a psychological disorder, or a biological disease, depending on the cause of the shutdown.

My point in articulating this example was to demonstrate that Behavioral Investment Theory offers a general conception of animal behavior that is organized in a way that leads to new formulations of important constructs in human psychology and provides new ways to achieve conceptual clarity on constructs that have historically been the source of much confusion. However, Behavioral Investment Theory is not just useful in clarifying concepts, but it is also useful in providing a frame that allows us to assimilate and integrate key insights from different theoretical proposals regarding how the mind (or behavioral investment system) works in general.

Behavioral Investment Theory and the Connection with Other Proposals

In the next section I review four separate works that each attempts to provide a unifying framework for understanding mind and behavior. Importantly, these perspectives are from various traditions in the field and are just a sample from which many other examples could have been chosen (e.g., Churchland, 1986; Gintis, 2009; Goodson, 2003), but nonetheless all are consistent with the broad conceptual framework provided by Behavioral Investment Theory. By reviewing them here, I hope to show more clearly key aspects of Behavioral Investment Theory and show why and how it consolidates a multitude of perspectives in a coherent manner.

W. T. Powers' Perceptual Control Theory

One of the aspects of cognitive science that I initially had trouble digesting was that it seemed to suggest that the mind is an input–output computational system that works via a series of sequentially arranged "if…then" commands. This model, of course, stemmed from the analogy that the mind is like a computer, and computers were generally programmed by such commands at the time. While I could see that some of my verbal–conceptual reasoning structures might function that way, it did not seem plausible that such "if. . .then" programming would effectively account for the incredibly dynamic aspects of behavioral engagement. And I wasn't clear on the relationship between such computational models and behavioral concepts like classical and operant conditioning, nor emotions such as joy or despair. Moreover, it was clear from neuroscience that the brain was not a sequential information processing system, but instead was a massive parallel information processing system. It wasn't until I discovered William Powers' (1973) perceptual control theory that I saw that these pieces could all clearly fit together.

Deeply concerned with how orthodox behaviorism had banished the concept of purpose from science, William Powers (1973), a humanistically oriented engineer, developed *perceptual control theory*, which provides a powerful model that accounts for the purposeful nature of animal behavior. In the opening chapter of his classic work, *Principles of Psychology*, William James (1890) gave a wonderful descriptive contrast between the behavior of physical objects and the behavior of animals. Specifically, he contrasted the behavior of bubbles in water to that of a frog and iron fillings attracted to a magnet to the behavior of Romeo and Juliet. James made the point that while we can imagine "forces" causing each set of behaviors, there nonetheless is a qualitative difference.

For the sake of argument, let's use a purposeful word "want" to describe the behavior of the iron fillings attracted to a magnet and the bubbles rising to the surface in water. By observing their behavior, we might initially say that the fillings "want" to reach the magnet and the bubbles "want" to reach the surface. But James pointed out that if an obstacle is placed in their path (e.g., a card is placed between the magnet and fillings or a piece of wood blocks the bubbles from reaching the surface) neither the fillings nor the bubbles will change their behavior to reach the supposed goal. James contrasted this state of affairs to the behavior of a frog and Romeo and Juliet.

> Unlike the bubbles, a frog will not perpetually press his nose against [the jar's] unyielding roof, but will restlessly explore the neighborhood until by re-descending again he has discovered a path around its brim to the goal of his desires. The case is the same for Romeo and Juliet: Romeo wants Juliet as the filings want the magnet; and if no obstacles intervene he moves toward her by as a straight as line as they. But Romeo and Juliet, if a wall be built between them, do not remain idiotically pressing their faces against its opposite sides like the magnet and the fillings with the card. Romeo soon finds a circuitous way, by scaling the wall or otherwise, of touching Juliet's lips directly. With the fillings, the path is fixed; whether it reaches the end depends on accidents. With the lover it is the end which is fixed, the path may be modified indefinitely. (James, 1890, p. 4; cited in Cziko, 2000)

James pointed out that we do not see purpose in immaterial events because the means or the forces are "fixed" (yet, for an analysis of when and why we humans do sometimes see purpose in such events, see Shaffer, 2008). In contrast, in animal behavior it is the ends that are fixed, while the means will vary indefinitely. This is the key feature that defines purposeful behavior. And William Powers realized that the control systems theory provided a framework to explain the purposeful nature of animal behavior.

Powers' control theory model is built on Wiener's (1961) notions about feedback. Both the cruise control in your car and the thermostat in your home are examples of simple control systems that operate on feedback. There are a minimum of three components to such a system: (1) an input sensor; (2) a reference goal; and (3) an output mechanism. In the case of cruise control, the speedometer is the input sensor, the speed at which you set the cruise control is the reference goal, and the output mechanism is the addition or removal of gas to accelerate or decelerate the car. Now put this system in motion. Say you set the cruise control reference level to 65 mph. If you start to head up a hill, your speed will drop and a discrepancy will emerge between the set reference level and the actual speed measured by the speedometer. That discrepancy will activate an output, more gas. If too much gas is added, and say your speed reaches 68 mph, then that difference results in decreasing the gas. In so doing, the car maintains a fairly specific speed, despite significant changes in external factors, such as wind and incline of the road.

Control theorists often use the following formulation to define the relationship between the key variables: input − reference goal => output. This formulation exists in contrast to the basic S→R formulation of behaviorism and the Input→Output formulation of cognitive science approaches. This control theory equation is called a negative feedback loop because the output is designed to reduce the discrepancy between the input and reference goal. Latter, I will adjust the control theory formulation slightly and characterize it as the P − M => E equation (or formulation), where the "P" stands for perception, the "M" for motivation, and the "E" for emotion.

Powers (1989) developed a detailed theoretical and empirical framework that applied the insights of control theory to animal and human behavior. One of his major epiphanies was that the control theory model not only accounted for purposeful behavior in a straightforward way, it also turned the traditional behavioral Stimulus→Response conception of causation on its head. Instead of sensory inputs directly causing behavior in a sequential and linear fashion, in Powers' model it was more appropriate to reverse the causal arrow and instead think of behavior controlling perception. That is, for Powers, animal behavioral output functions to reduce the discrepancy between the perceived actual state and the reference desired state. Thus, the function of the behavior is the control of perception, rather than the stimulus causing the behavior. For example, in traditional stimulus response theory, the crossing double yellow lines on the road is associated with punishment, thus as your car nears them, you turn back the other way out of avoidance. The double yellow lines are seen as a stimulus and turning back is the resultant response. In contrast, in control theory language, your mind has a reference goal state of the relationship between the car and the lines on the road and you work to maintain a minimal discrepancy between the perceived state and that reference goal state. In the former formulation, the stimulus triggers the response; in the latter, the response is a function of attempting to "control" your perceptions. The differences are seen in the way the two perspectives describe such variables. Traditional behaviorists speak of controlling variables (i.e., variables that control the animals behavior), whereas control theorists speak of controlled variables (i.e., variables that the animal attempts to control). The distinction is crucial because of the philosophical implications. In the former model, humans are mechanistic, like billiard balls. In the latter, they are purposeful, agentic creatures.

Despite the important philosophical differences between traditional behaviorism and Power's control theory, the perspectives are really only incompatible at the extremes of the continuum. Consider that although Powers developed his system in direct response to behavioral philosophy, it turns out that during the same period an integrative behavioral theorist was developing a model that cut across various learning theories and approaches and created a framework that actually is very consistent with Powers' control theory model.

Arthur Staats' Psychological Behavioral Theory of Learning

As was mentioned in Chapter 2, Arthur Staats saw the tremendous problems associated with psychology's fragmentation, and he devoted his career to developing a unified theoretical and methodological system that he argued could stretch from the simplest of animal behaviors to the most complex human thought patterns (Staats, 1996). Staats sought unification through a system that would first mend factions within behavioral theory (e.g., Hullian, Skinnerian, Neo-behavioral perspectives) and then, from that foundation, build conceptual and methodological bridges to traditional psychology (e.g., developmental, personality, social). Because of this bridge building between psychology and behaviorism, he called his framework psychological behaviorism.

Staats' argued that psychological phenomena consisted of multi-leveled processes, and we needed to start at the bottom and work our way up. The foundation of his model is a basic theory of learning that emphasizes the evolutionary function of emotional responses. Specifically, in Staats' model it is the emotional response mechanisms that allow an animal to learn.

> It is the emotion–reinforcer relationship. . .that has been built into organisms through biological (evolutionary) development. . .Emotional responses are elicited by stimuli that are biologically important to the organism—either to obtain, like food, or to avoid, like painful stimuli. . .That is the essential behavioral reason why emotions are important, because they define what will be reinforcing for the organism in the sense of affecting what behaviors the organism will acquire. (Staats, 1996, p. 41)

Always looking for connections between disparate view points, it is useful to notice the similarity to this formulation with the quotation I offered earlier from the evolutionary biologist Richard Dawkins in the principle of learning section. Both are arguing that instead of coming equipped with a set genetically preprogrammed list of commands of how to act, nature has built into animals' emotional systems that assess whether the animal is effectively moving toward or away from its goals and shift the behavioral output accordingly.

In a nutshell, both are also arguing for the pleasure pain parallel fitness principle mentioned earlier. Pleasure signals animal–environment relationships that were associated with enhanced fitness, whereas pain signals animal–environment relationships that were associated with diminished fitness. It is hardly accidental that we generally feel pleasure when eating, but pain when being eaten! This principle includes the proposition that there are two broad categories of emotional responses (positive and negative) and two broad categories of behavioral responses (seeking/approaching and avoiding/withdrawing). The fundamental reason that there are two broad categories of behavioral responses is that there were two broad kinds of evolutionary problems animals needed to solve, namely acquiring benefits and avoiding losses. For those more quantitatively inclined, it is useful to note that one can think about this in basic statistical terms. Some animal–environment relationships positively covaried (or correlated) with survival and reproductive success, whereas others negatively covaried with survival and reproductive success.

Staats (1996) called his model a three function learning theory because he asserted that the emotional response performs three crucial functions associated with learning and motivation. First, the emotional response functions to produce classical conditioning. Classical conditioning is perhaps the most basic learning process and refers to the manner in which neutral stimuli can come to elicit a response after being paired with an unconditioned stimulus. Staats' point was that unconditioned stimuli are by definition stimuli that elicit an emotional response, and if no emotional response occurs, then no pairing or learning occurs.

Second, Staats' argued that the emotional response functions to produce operant conditioning because when an emotional response occurs after a behavior, it will reinforce the behavioral response. That is, activities that produce an outcome associated with the increase in positive emotions or decrease in negative emotions will be strengthened, whereas activities that result in a decrease in positive emotions or an increase in negative emotions will be extinguished.

Third, emotional responses function as incentives. When a stimulus elicits a positive emotional response, the animal will approach the stimulus, and when it elicits a negative emotional response the animal will then emit avoidance or escape behaviors.

Although closely related, the primary difference between the incentive and reinforcement functions for Staats is in terms of the timing between the stimulus and response. The emotional response is serving as an incentive function when the stimulus is presented first, but a reinforcing function when the emotional response occurs after the actions. For example, when a child sees a lollipop and has a positive emotional response based on past experiences eating a lollipop, the anticipated pleasure is serving an incentive function. When the child is struggling to open the wrapper and finally rips it with his teeth, shifting his experience from frustration to pleasure, then the emotional response has reinforced the behavior and he is more likely to rip it with his teeth in the future. Staats (1996) has demonstrated that the three function learning model can incorporate central findings from behavioral research on phenomena such as extinction, generalization, discrimination, and intermittent conditioning. He also argued that this model can form a base for understanding higher cognitive processes in humans.

Although schisms abound in the mind and behavioral sciences, a major one is between behaviorism and evolutionary psychology. The former emphasizes general changes in behavior that are the result of experience and consequences, whereas the latter tends to emphasize specific information-processing modules built by natural selection. Moreover, evolutionary psychologists often seem to dismiss or argue against domain general processes, such as Skinner's behavioral selection. With that in mind, it is interesting to note that a student of evolutionary psychologists' John Tooby and Leda Cosmides developed an evolutionary approach to animal behavior that is very consistent with Staats' learning theory.

Peggy LaCerra's Adaptive Representational Networks

Earlier in the chapter I described why LaCerra's (2003; LaCerra & Bingham, 2002) conception of how intelligence systems manage the expenditure of behavioral energy is directly consistent with Behavioral Investment Theory, especially the first principle of energy economics. In delving into the mechanisms of how intelligence systems are constructed, LaCerra introduced the concept of Adaptive Representational Networks (ARNs). ARNs are the fundamental building blocks of intelligence systems that LaCerra argued function to encode important life events by linking together four crucial components: (1) the internal state of the animal (i.e., how it is feeling inside, cold, hungry, etc.); (2) the sensory features in the environment (i.e., what it is perceiving to be going externally); (3) the behavioral response (i.e., what the animal did in that situation); and (4) the adaptive value of that response (i.e., whether the outcome was good or bad). For example, if you felt hungry, saw hotdog, ate it, and then felt sick later, that episode would come to constitute an ARN. Because in this case the adaptive value of the response was negative, next time you were hungry and saw a hotdog, it is likely you would continue searching for something else to eat.

LaCerra argued that ARNs build up over time—scene by scene, episode by episode—to form a catalogue of memories that link together internal states, external conditions, actions, and the subsequent value of the consequences. This catalogue of memories forms the bank of information that we draw on to make predictions about which behavioral investments will be most effective in the current situation. The dynamic interrelation between these four components and the manner in which they regulate actions via negative feedback loops is directly congruent with Powers' model. In both models, input is referenced against some desired goal state and the discrepancy activates a behavioral response. LaCerra and Bingham add the point that the outcomes of those actions are then stored and used to guide future behaviors, a point Powers would undoubtedly agree with.

David Geary's Motive to Control as the Central Principle of Mind

David Geary is an evolutionarily minded cognitive and educational psychologist who has argued that the motive to control the flow of resources is the fundamental organizing principle of mental systems. Hopefully, this should sound quite familiar.

Indeed, Geary and I had been working independently on our various proposals when we met online and discovered the striking similarities. He authored a paper detailing his model and drawing out the parallels between it and Behavioral Investment Theory. In it he wrote

> My proposal is that the brain and mind of all species has evolved to attend to and process the forms of information, such as the movement patterns of prey species, that covaried with survival and reproductive outcomes during the species evolutionary history. These systems bias implicit decision-making processes and behavioral responses in ways that allow the animal to attempt to achieve access to and control of these outcomes, as in prey capture. . .The thesis is here. . .that the human motive to control is indeed an evolved disposition and is implicitly focused on attempts to control social relationships and the behavior of other people, and to control the biological and physical resources that have historically covaried with survival and reproductive prospects in the local ecology. (Geary, 2005, pp. 23–24)

Geary argued for four different components of control-related mechanisms: (1) affective; (2) cognitive; (3) conscious–psychological; and (4) behavioral, and each system is embedded in various domains of the brain and nervous system. The behavioral systems are the most basic and perform the procedural sensory-motor outputs, such as those involved in walking or catching a ball. In a manner closely related to Staats' formulation, Geary argued that affective systems guide the behavioral strategies, and provide feedback to the individual regarding the effectiveness of the behavioral strategies. Specifically, positive feelings provide reinforcement when strategies result in the achievement of significant goals, and negative feelings result when behaviors do not result in such ends. Geary used the more narrow definition of the term cognitive, meaning executive functioning and working memory that allow individuals to mentally represent and manipulate information processed by sensory and perceptual systems, as well as inhibit behavioral impulses predicted to yield problematic outcomes. Conscious–psychological systems offer an explicit representation of situations of significance. Geary described the functioning of these mechanisms in the language of control theory. That is, he views these systems as generating internal representations of the way the world is desired to be which are contrasted to the way the world actually is, and behavioral strategies are invoked to reduce the discrepancy.

One of the real strengths of Geary's model is how it provides an understanding of the classic nature–nurture divide. He offered what might be considered a middle-ground position between the evolutionary psychologists and traditional learning theories with an idea he calls soft modularity. As mentioned above, the evolutionary psychologists tend to view the mind as made up of domain-specific models which solve specific adaptive problems (e.g., we have specific cheater-detection modules or mate-guarding mechanisms), in contrast to learning theorists who tend to emphasize general processes by which animals adapt to their environment (e.g., operant and classical conditioning processes). Although as Geary and Huffman (2002) point out, "all serious theorists now agree that there are inherent gene-driven constraints on and experience-based—especially during the developmental period—modifications of brain organization and cognitive functioning" (p. 667), there is much debate on the relative amount of genetic constraint versus experiential plasticity in various domains of mental functioning.

An advantage of Geary's system is that it provides a helpful way of understanding and framing the way that evolution has prepared animals to learn. Stemming from his formulation that the foundational base of the mind is the motive to control resource flow, Geary makes the following assertion about when we would expect to see high levels of constraint and high levels of plasticity:

> [Important] resources generate information patterns that range from invariant to variant. Invariant information is consistent across the generations and within lifetimes (e.g., the prototypical shape of a human face) and is associated with. . . processing of information that is implicit and results in bottom–up behavioral responses. Variant information varies across the generations and within lifetimes (e.g., as in social dynamics) and is associated with plastic brain and cognitive systems and explicit, consciously driven top–down behavioral responses. (Geary, 2005, p. 21)

In other words, there are some informational patterns that remain constant over long periods of time (i.e., the generations) and other information patterns that vary. The former will be hard-wired into the programming of the nervous system, and animals will have relatively automatic response patterns when processing information that was consistent across the generations. In contrast, learning mechanisms will be in place to allow the animal to adjust in response to the information patterns that are much more variable and situation dependent. Soft modularity, then, is the idea that evolution will build in to the information processing architecture basic ways of framing representations and responses, but experience and associated learning mechanisms will then fill in these frames in detail so that the animal can adjust to the idiographic nature of its particular environment.

The Influence Matrix, the subject of Chapter 4, is very consistent with Geary's motive to control formulation and is a clear example of soft modularity. The Influence Matrix posits that there are certain aspects of social relating that were relatively constant across the generations. For example, competition and cooperation are two social exchange processes that all social animals have to navigate, and thus evolutionary forces have shaped the human social motivational system with distinct information processing systems for framing these two types of relational processes. And yet, who we should cooperate or compete with and when is of course completely dependent on the specific situational context. Thus, evolution has not built us with pre-wired commands on when to compete or cooperate but instead has fashioned mechanisms that evolve with experience to guide us in the ever changing social world in which we currently live.

If this review of other proposals has been successful, a basic outline of how animals calculate their behavioral investments should be emerging. Each of the various proposals outlines mechanisms built by evolutionary processes, which include inputs that are evaluated as positive or negative resulting in responses that are evaluated again, creating a behavioral feedback loop. In the next section, I offer a schematic of the architecture of the human mind and specify the nature of the behavioral control feedback loop in what I call the P − M => E formulation.

Behavioral Investment Theory and the Architecture of the Human Mind

Open any standard introductory psychology textbook, and you are likely to see in the first chapter or two an overview of the various paradigms and perspectives in psychology (e.g., behaviorism, social-cognitive, psychodynamic, cultural, evolutionary, etc.) followed by separate chapters on biological bases of behavior, sensation and perception, motivation, emotion, cognition, memory, learning, and consciousness, in addition to overviews of aspects of the field, such as clinical psychology and various approaches to psychotherapy. Although there are clear advantages to this format in terms of articulating lines of research and major empirical findings, a problem arises because not only are the theoretical paradigms often defined against one another, but the various psychological domains and lines of research are also not interconnected in a way that gives students a feel for how the human psychological system functions as a whole. When you look at the way material is presented and ask questions such as, “How does evolution relate to motivation?” or “How do perception and emotion relate to learning?” or “How do operant principles relate to cognitive phenomena?” the answers are usually not forthcoming.

In contrast to this state of affairs and in accordance with the overarching goal of this book, this section offers a schematic of the human mind that attempts to make explicit the interconnections between major psychological phenomena so that one can obtain a feel for how the behavioral investment system works as a whole (Fig. 3.2). Although like all maps this schematic emphasizes certain features and

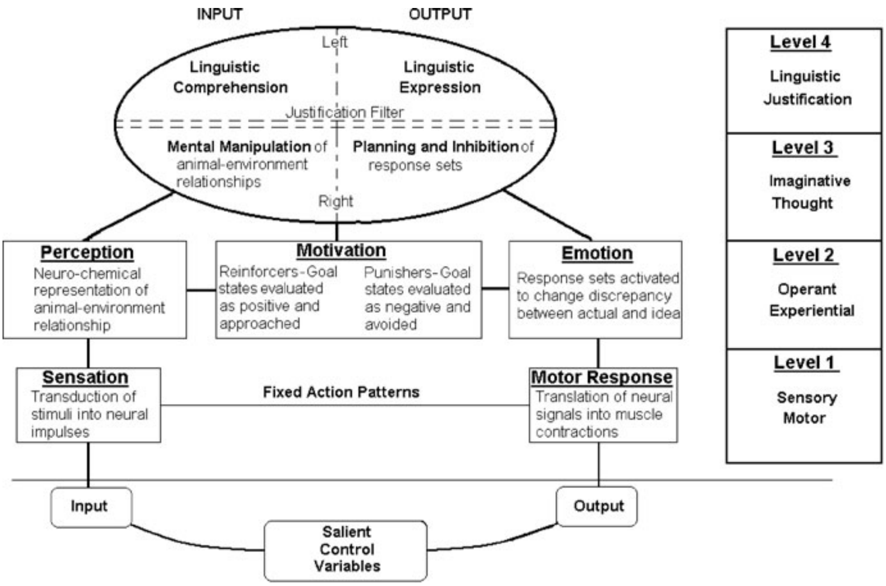

Fig. 3.2 The architecture of human mind

leaves out others, the value in this model is that it provides a clear framework for understanding how the human mind functions to compute behavioral investments in a manner that connects evolutionary, neurological, computational, and behavioral systems approaches.

Before proceeding to clarify the components of the diagram, I need to offer a few words about the definitions of important terms, including mind, brain, cognition, and consciousness. For starters, let me differentiate Mind (upper case "M" from mind). When capitalized, Mind refers to the third dimension of complexity on the ToK System and consists of the set of mental behaviors. Mental behaviors, which are discussed in more detail in Chapter 7, are behaviors of the animal as-a-whole mediated by the nervous system that produces a functional effect on the animal–environment relationship. Hunting, mating, defending a territory are commonplace examples. But mental behaviors can also be covert. Perceptions, feelings, imaginings, and even nonconscious cognitive processes are also considered mental behaviors. In slight contrast to Mind, when I am referring to the "mind" as in the aforementioned diagram, I am referring to the architecture of the neuro-information processing system, as well as information instantiated in and processed by that system. In this sense, the mind is the neuro-computational control system that allows the animal to behave as a coordinated whole.

The brain is the biophysical material that mediates covert and overt animal behaviors. To consider the difference between the mind and the brain, think about a book. On the one hand, there is the material dimension of the book, which consists of its weight, and the structure of the molecules that make up the ink and pages. This is akin to the brain. On the other hand, there is the informational content of the book, which is akin to the mind.

The two other terms that need to be defined are cognition and consciousness, which are two kinds of covert mental behaviors. Let's start with cognition. Cognition is a term with at least two different meanings in the literature. Sometimes the term is used in the broad sense to refer to general neural-information processing. When used this way, the broad definition of cognition includes perceptions, imaginations, and language, but also motivations, emotions, and unconscious processes. The broad definition of cognition is essentially equivalent to the definition of the mind discussed above. On the other hand, the term cognitive is frequently used in a more narrow sense of higher thought processes (commonly, language-based thinking), and such processes are conceptualized as separate from motivation and emotion. Some researchers in human cognition (e.g., reasoning and intellect) and those who do cognitive psychotherapy tend to emphasize this meaning.

Consciousness refers to the subjective, first person phenomenological world. It is the experienced world of senses, images, feelings, and conscious thoughts. According to the unified theory, consciousness is a particular kind of cognitive process; thus, it is a form of neuro-information processing that through some as-yet-to-be-determined process becomes experienced. There are two broad domains of consciousness: one of which is often referred to as sentience and is experiential and includes our perceptions, feeling states, and nonverbal images; and the other self-consciousness consists of language-based self-reflective thought.

Figure 3.2 is a model of cognitive processing (broadly defined) that divides the human mind into four levels of neuro-information processing: (1) Sensory-Motor; (2) Operant Experiential; (3) Imaginative Thought; and (4) Linguistic Justification. Below I briefly describe each level and related features in the diagram. It is, however, important to note that although there is much heuristic value in dividing the human mental system into levels, it is also the case that there are constant interactions between the levels in very complicated feedback loops so that it is important not to think of these levels as completely separated from one another.

Level 1: Sensory-Motor

The sensory-motor level is the most basic form of neuro-information processing and the kind that first evolved. It is called sensory motor because there is a fairly immediate connection between stimulus and response. Sensory-motor processes are the most "hard wired," and fixed in the sense that the animal does not require much experience to develop them. There are some animals, like jellyfish, snails, and worms, whose nervous systems likely function completely at the sensory-motor level. In humans, the basic reflexes, such as blinking your eye in response to a puff of air or pulling your hand away from a hot stove, are examples of sensory-motor processing. Despite its simplicity, basic forms of learning, such as habituation and sensitization, take place at the sensory-motor level. Habituation is a decrease in a reflex response resulting from repeated presentation of an initiating stimulus. It is arguably the most basic form of learning and there are functional parallels to this form of learning present even in single celled organisms, which obviously lack a nervous system completely. Sensitization is essentially the opposite, and refers to the process by which an animal learns to increase its reflexive responses to noxious or novel stimuli. Although instinctual reflexes exist at the sensory-motor level, so do acquired habitual motor patterns like walking, which are automatic and generally nonconscious.

The highest forms of sensory-motor behavior patterns—patterns that, in terms of complexity, exist just beneath operant behaviors—are fixed action patterns (FAPS). FAPS are well-defined motor patterns that are "fixed" because they emerge without much trial and error learning, are triggered by particular "releasing stimuli", and are species typical. Historically considered instincts, FAPS are elaborated networks of reflexes that function to coordinate movement into much more complicated sequences than simple muscle contractions involved in, say, the human blink reflex. FAPs were studied extensively by the founders of ethology, Niko Tinbergen and Konrad Lorenz. A classic example of a FAP is the begging behavior of young herring gulls, which is "released" by a red dot on the parent gull's beak (Tinbergen & Perdeck, 1950). Although humans generally do not exhibit complicated FAPs that are present at birth, there are some, such as the rooting reflex.

The difference between sensory-motor processing and higher forms of thought was on striking display for me and my wife during a trip to the emergency room with our son Jon in the fall of 2009. He was playing soccer, and he tripped and landed awkwardly, breaking both bones in his right arm. After the x-ray, the medical team came in to set the bone. The doctor asked us if we would feel more comfortable waiting outside. They informed us that although they were going to anesthetize him, it was still the case that "the body reacts to the process." We decided to stay and observed them put him under and then we watched as they worked hard to set the bone and he moaned and his body writhed. After the anesthetic wore off, he quickly returned to a full state of consciousness with no recollection or awareness of what had occurred. Nevertheless, his mental system had been operating at the sensory-motor level.

Level 2: Operant Experiential

Operant behavior patterns are much more complex, fluid, and plastic than are sensory-motor behaviors. Whereas sensory-motor behaviors are characterized by relatively rigid, immediate, and reactive behaviors, operant behavior patterns are more proactive and characterized by dynamic seek-and-approach or avoid-and-withdraw sequences. Skinner referred to operant behaviors as "voluntary" and contrasted them with "respondent" behaviors that are more reflexive in nature (i.e., level 1 processes). Operant behaviors appear later in the evolutionary sequence than sensory-motor behaviors, perhaps first emerging with the jawed fish. Fish, amphibians, reptiles, birds, and mammals all demonstrate operant behavior patterns.[3]

Operant behavior is characterized in the current formulation as operating on the P − M => E equation, where the "P" refers to perception, "M" to motivation, and "E" to emotion. Verbally, this formulation translates into *perception of an actual state relative to a motivated state leads to an emotional state*. This is a control theory formulation, and I am hopeful that the earlier reviews of other proposals will make this seem rather straightforward. Recall that the basic structure of a control theory equation is input − reference goal = output. Thus, "P" refers to the perceptual input. Perception is a higher level mental process than sensation and occurs via the integration of sensory inputs that result in a meaningful representation of an object or event. Perception is a consequence of both bottom–up processing, which refers to the pattern of sensory inputs, and top–down processing, which refers to the individual's knowledge, memory, and expectations. (Of course, the degree to which top–down input influences the perception depends tremendously on the cognitive complexity of the animal). Researchers have analyzed these processes extensively, and there are many good books on the principles and mechanisms of both bottom–up and top–down perceptual processes (e.g., Shipley & Zacks, 2008).

[3] Interestingly, so do many insects, which is somewhat surprising given that their brains are smaller than the head of a pin. The complexity of insect behavior patterns deserves close attention from psychologists, for they may well challenge some of our most basic assumptions about brain and mind. Nevertheless, because our evolutionary history is sufficiently divergent from such creatures, I will not delve deeply into the latest research on insect behavior (see Prete, 2004).

To give an example illustrating the influence of top–down expectancies, when I was in high school, a friend played a joke on me by switching the coca-cola in his can for cream soda. He then told me the coke tasted weird and asked me to take a sip. I did and immediately spit it out. It tasted horribly! He then revealed his trickery, which was particularly interesting to me because I like cream soda. Under normal circumstances when the bottom–up sensory inputs line up with my templates for cream soda, I recognize it, and it tastes good. On the other hand, when there was a mismatch between top–down expectations and bottom–up sensory input, the perceptual experience was radically different, and I reacted with disgust.

Motivation, the "M" in the equation, refers to valued goal states that the animal is working toward attaining or avoiding. Stated slightly differently, the motivational system has templates for benefits to be approached and costs to be avoided. There are two broad classes of goal states, approach states and avoidance states, because the basic templates emerged as a consequence of evolutionary processes, and there were animal–environment relationships that either positively or negatively correlated with survival and reproductive success. Much research has demonstrated that there are two broad behavioral systems, which are often referred to as the behavioral activation and behavioral inhibition systems that respectively activate seek-and-approach or avoid-and-withdraw behavioral patterns.

In accordance with Geary's soft modularity, evolutionary processes built the basic frames for the motivational templates, and then learning and experience fill them in, greatly elaborating on them. For example, it is clear that many birds and mammals at birth have motivational templates for maintaining proximity to their parents, but these animals also learn via experience to identify their particular parents. Likewise, the Influence Matrix posits that humans come equipped with motivational templates to approach high social influence and avoid the loss of influence. However, it is experience that dictates who the important individuals to influence are and the methods by which such influence is achieved.

Motivational templates become activated depending on both the perceptions and the internal state of the animal. For example, imagine you are camping and after about 6 hours between meals an internal state of hunger activates a goal template of eating a sandwich in the cooler. In that state you would be motivated to reduce the discrepancy between where you were and the end goal state of eating the sandwich. However, if you stepped out of your tent on the way to the cooler and saw a bear, a very different goal template would be activated, which would be one of your being safe, far away from the bear. Indeed, the perception of a bear would initiate a cascade of bodily responses, readying you for fight or flight. Some of those responses would include a massive sympathetic nervous system shift activating muscles and attention and energy would be diverted away from processes like digestion, and it is likely you would not feel hungry for quite some time after that occurrence.

Although separable, perceptions and motivations are intimately intertwined. It is rather obvious that perceiving something can activate a motivational state. But the reverse is also true in that your motivational state will greatly impact your perceptual experience. Consider, for example, how you perceive food before as opposed to after eating, or sexual activity before and after sex. Along these lines, research on the Rorschach Inkblot Test has demonstrated that hungry individuals will see more foods and food-related activities in the inkblots than sated individuals (Epstein, 1961). Similarly, Bruner and Goodman (1947) found that children from impoverished backgrounds perceived a quarter as literally larger than children from wealthy backgrounds. More recent work has demonstrated that individuals who are fatigued will perceive distances that they need to walk as being longer than individuals who are rested (Proffitt, Stefanucci, Banton, & Epstein, 2003).

Emotions, the “E” in the equation, organize the animal’s response set and are activated based on the relationship between perceptions and motivations. The two broad domains of satisfying and aversive affective reactions form the foundational base of the emotional response sets. The reduction of a discrepancy between perceptions and an approach goal state activates a positive emotional state (e.g., satisfaction, joy). Increasing the discrepancy between perceptions and an aversive state also activates positive affect, although of a slightly different tenor (e.g., relief, relaxation). In contrast, decreasing the discrepancy between perceptions and an aversive state results in negative emotions (e.g., fear, hurt), and so does increasing the discrepancy between perceptions and an approach state although again of a slightly different tenor (e.g., frustration).

The strength of the emotional response is tied to the nature of the motivational state. Harking back to Maslow’s hierarchy of needs, significant and immediate threats to the physiological needs or safety (or such needs in loved ones) are likely to activate the strongest emotional responses. Thirty seconds without air and all resources are likely to be devoted to moving toward an oxygen source. It should also be noted that emotional responses are tethered to the rate of change between perceptions and goal states. Anyone who has experienced frustration at not making enough progress knows that one can be moving toward a goal, but if the rate is not fast enough, it will produce a negative emotional response. Whereas sensory-motor processing takes place in individual neural nets, the spinal cord and brain stem, the perceptual-motivational-affective processes associated with operant behavior patterns is associated with structures in the limbic system.

Level 3: Imaginative Thought

Imaginative thought takes place in the cortex and refers to the ability to manipulate mental representations into simulations of behavioral investment patterns and then be guided by anticipated outcomes. The classic demonstration of higher nonverbal thought in animals was Wolfgang Kohler’s work on insight in chimpanzees. Kohler placed fruit that was just out of reach of the chimpanzees. He then provided materials that the chimps could use to achieve their goal. For instance, he placed boxes that the chimps could stack and then climb on to reach fruit, and placed sticks that could be stuck together to reach fruit outside the cage. These animals clearly did not engage in overt trial and error actions. Instead, their behaviors strongly suggested they achieved insight via the mental manipulation of variables that allowed them to simulate actions and changes until a clear pathway to the goal could be foreseen.

Why, from a Behavioral Investment Theory perspective, would such capacities have evolved? From a functional perspective, the answer is straightforward. Assuming the mental simulations have at least some predictive validity, they are valuable because they are far less costly to run than actually engaging in the behaviors overtly. Imaginative thought occurs in the neocortex, and it is not accidental that mammals are generally seen as engaging in more complex thought than reptiles or birds and have larger portions of their brains being made up of the neocortex. Likewise, primates are generally considered to be the cleverest of mammals and have proportionally the most neocortex. The neocortex in humans is huge.

Level 4: Linguistic Justification

Level four in the diagram represents the intersection of language, culture, and self-consciousness. Although some other animals are capable of complex nonverbal thought, only humans come equipped with the capacity to generate a symbolic-syntactical representational system (Penn, Holyoak, & Ponvelli, 2008). This system is referred to by some as the language acquisition device (Pinker, 1994), although there remains significant debate about its exact nature (see Greenspan & Shanker, 2004). At a very basic level, this system allows humans to symbolically label perceived objects and their transformations in time in the form of nouns and verbs and differences between things in the form of adjectives. As is evidence by the enormous variety of the world's languages, the system is an open system; children learn the language of the culture they are born into. Chapter 5 focuses on the dimension of linguistic justification via the lens of the Justification Hypothesis.

The final aspect of the diagram that at least needs to be noted is the section labeled salient control variables. In a nutshell, the salient control variables are those aspects of the environment that influence the behavior of the animal. These consist of the set of variables and processes studied by traditional behavioral science research. I label them control variables in a nod to both behaviorists, who see these variables as the environmental determinants of behavior, and to control theorists who emphasize the notion that there are certain variables that the animal is attempting to control.

An Everyday Example

As I mentioned in the preface, qualitative generalizability refers to sharing information that is intuitively understandable by most people and is one of the principles guiding the construction of this book. In accordance with that goal and in order to provide a fuller sense of how these layers might operate in everyday life, I will offer the following commonplace situation and then break it down according to the diagram to give a feel for how to apply it.

> Jason, a nine year old boy, is playing a video game, when he realizes he is feeling a little hungry. On the way to the kitchen to get an apple, he spies some brownies in the pantry. His mother is upstairs and one of the rules of the house is no sweets before dinner. Overtly, he stands at the pantry for several seconds, staring at the brownies before finally walking

> away, grabbing an apple, and returning to his video game. While he was standing there, he had the image first of opening the package and wolfing down several brownies. However, that was followed by an image of his mother discovering the open package and punishing him. He then imagined asking her permission but remembered her answer would likely be an irritated "No"! He then thought about hiding the empty package and about lying, saying his sister ate them, but both images made him feel guilty. He then told himself it was not right to eat dessert before dinner and that the apples his mother had just bought were the kind he liked.

Starting at the first level, there is the basic sensory input that Jason is receiving. This refers to the various physiochemical changes that are taking place both inside and outside his body and are being translated into neuronal messages. On the output end, efferent neurons are translating neuronal messages into coordinated muscle movements, allowing him to play his game and then get up and walk over to the pantry. Shifting up a level, the sensory inputs are organized into perceptions, which are higher level integrations of sensations into meaningful representations. Thus Jason perceives (rather than senses) the brownies, meaning that he integrates the various sensory aspects of the shape, color, location of the brownies into a meaningful category or concept. On the same level as perceptions, motivations are the goal states that organize and direct action and attention. A goal state is a schema or template of an animal–environment relationship to be sought and approached or avoided and withdrawn from. In Jason's case, the perception of brownies was referenced against the experience of hunger and memories of eating such brownies in the past (think here of La Cerra's Adaptive Representational Networks). This led to the image of him eating brownies, which can be thought of as a reference goal state (think here of Power's control theory model). The image of him eating the brownies was initially associated with a pleasurable emotional response, which functioned as an incentive to approach the brownies and reduce the discrepancy between where he was and the reference goal state (think here of Staats's three function theory).

But Jason did not eat the brownies. So what happened? To understand this, we need to move up a level on the diagram. Rather than simply responding to present perceptions, nature has equipped many animals (especially mammals) with the capacity to simulate outcomes without actually having to go through them, an ability many would refer to as thinking. So we see in the example that Jason simulated several possible behavioral investment pathways. To do so, he needed to first inhibit the actual act of grabbing the brownies while he ran through a sequence of events that ended with him being punished by his mother, or feeling guilty because he had disobeyed. The images of those consequences were aversive, thus he experienced an incentive to adopt a different path.

I have seen my dog wander by the pantry where we keep our treats, stop, glance back at me, and then walk away. There have been other times when I have come down stairs only to find that she had gotten into the goodies. I would argue that the analysis of these three levels of information processing would be relevant and indeed necessary to develop a holistic explanatory account of my dog's behavior. That is, at a minimum, animals like dogs engage in sensory motor, operant, and imaginative thought. But Jason engaged also in a fourth level, one that is not present in other mammals. He engaged in self-talk and framed his actions in terms of justifiability.

He did not just subconsciously associate punishment with eating the brownies. He knew there was a reason for not eating them. He considered attempting to justify why there should be an exception to the rules in this case to his mother, but decided she would not accept his justification. He then generated other possible pathways that would avoid his mother's disapproval, but decided these actions were unjustifiable. Then to help channel his behavioral investment pathway in a justifiable manner, he strengthened his own belief that eating dessert before dinner was not legitimate and justified to himself why eating apples was positive. Chapter 5 articulates the key elements of the justification system. The overall point here is that the schematic provides a basic conceptual framework for the key components of the human mind, or for those more behaviorally inclined, the behavioral investment system.

It is useful to conclude this section by noting that the idea that the mind–brain consists of different levels has a long and rather robust history. Perhaps the most notable layered view of the mind is Freud's structural theory of the id, ego, and superego and related topographical model of consciousness, preconsciousness, and unconsciousness. I hope that those familiar with Piaget's work on sensory-motor, preoperational, concrete operational, and formal operational stages of cognitive development will also see clear parallels and similarities with this model. The renowned neuroscientist, Paul MacClean, who was the former director of the Laboratory of Brain and Behavior at the National Institute of Mental Health, developed a layered model of the brain based on its evolutionary development. He called it the "triune brain theory" because he argued that the human brain is actually three brains in one, each of which was established successively in response to evolutionary need. The three layers are (1) the reptilian system, or *R-complex*, which consists of the brain stem and the cerebellum and functions to regulate activities associated with physical survival and maintenance of the body; (2) the limbic system, which consists of a network of related brain structures such as the hippocampus, amygdala, and houses the primary centers of emotion; and (3) the neocortex, which is the most recently developed outer portion of the brain that functions to allow for planning and prediction and other forms of higher thought (Maclean, 1993).

The Russian psychologist A. N. Leont'ev (1981) wrote extensively on the evolution of the brain and mind. He also argued for a four-layered view of mental processes that included: (1) an elementary sensory stage characterized by animals that react automatically and immediately to physical stimuli (e.g., worms and jellyfish); (2) a perceptive stage, which is characterized by animals that react to their perceptions of things (e.g., birds and mammals); (3) an intellectual stage (in "higher" mammals, such as apes), which is characterized by the capacity to mentally manipulate relations between objects in the environment; and finally (4) human consciousness, which is characterized by self-reflection and the formation of conscious judgments. Although I developed the Architecture of the Human Mind prior to learning about Leont'ev's work, the parallels are quite striking.

Marvin Minsky is a well-known professor at MIT who authored the influential book *The Society of Mind*, which likened the human mind to a society that had different information processing agents that functioned like different roles in society. In his most recent book, *The Emotion Machine*, Minsky (2006) updated his view of the mind. One of the most salient aspects of his argument is that the human mind is layered into six levels that have obvious parallels with the architecture of the human mind proposed here. The most basic level consists of instinctive reactions, which Minsky characterizes as built in *If→Do* reaction rules. The next level is learned reactions, which refer to the way animals develop new ways to react to situations based on lived experience. The next higher level deliberative thinking refers to planning, considering alternatives, and simulating possibilities. The fourth level in Minsky's scheme is reflective thinking, which refers to thinking about thinking. The fifth level, self-reflective thinking not only considers recent thoughts, but it also thinks about the entity doing the thinking. Finally, self-conscious reflection refers to thinking about how one ought to be in terms of higher ideals. Although Minsky argues for six levels instead of four, there are obvious parallels with the Architecture of the Human Mind diagram. Moreover, as Minsky (2006) himself notes the boundaries at the higher levels are "indistinct" (p. 147).

Conclusion

Behavioral Investment Theory is the joint point between Life and Mind on the Tree of Knowledge System and attempts to provide a comprehensive approach to animal behavioral science and the evolution of the animal mind by characterizing the nervous system as a computational control system that calculates the expenditure of behavioral investment on a cost-benefit ratio shaped by evolution and experience. Incorporated within Behavioral Investment Theory are six fundamental principles of animal behavior, which include the principles of: (1) energy economics; (2) evolution; (3) genetics; (4) neuro-computational control; (5) learning; and (6) development. To demonstrate its utility, the lens of Behavioral Investment Theory was applied to an important construct in human psychology, depression. It was argued that depression could be defined as a state of behavioral shutdown, and this conception reconciles seemingly conflicting perspectives of whether human depressive states are normal reactions, psychological disorders, or mental diseases.

A number of different broad proposals for mental functioning were then reviewed, including proposals offered by an evolutionary neuroscientist, an integrative behavioral learning theorist, a humanistically oriented engineer, and a cognitive psychologist, and showed the basic structure of mental process could be understood not as basic S→R formulations of traditional behaviorism, nor even Input→Output notions of early cognitive perspectives, but instead requires a control theory frame of P − M => E. The chapter concluded by broadly characterizing the different levels of mental processing that have evolved over the eons, culminating in humans with a language-based justification system. In the next chapter, Behavioral Investment Theory is utilized as the foundation to construct an integrative model of human social motivation and emotion called the Influence Matrix.