

Human Nature in a Nutshell

A unified approach to psychopathology would be impossible without an integrated working model of the mind and its functions. In Part I of the book (Chapters 1–4), I draw on evolutionary psychology, anthropology, and neuroscience to sketch such a model. This initial chapter has two goals—first, to present a brief overview of human behavior and development from an evolutionary perspective and, second, to introduce a toolkit of biological concepts and theories that provide the necessary background for the remainder of the book.

WHO WE ARE

Homo sapiens is an ape whose lineage separated from that of chimpanzees and bonobos 5–7 millions of years ago and acquired its modern anatomical features in Africa about 300,000 years ago (Hublin et al., 2017; Jobling et al., 2004). Until the invention of agriculture at the end of the latest glacial period—a mere 12,000 years ago—human subsistence worldwide was based on hunting, fishing, and gathering. The use of sophisticated foraging techniques that make extensive use of manufactured tools is part of the *human adaptive complex*—a suite of coevolved traits that contribute to define the ecology of our species (Kaplan et al., 2000). The human adaptive complex includes the exploitation of high-quality food sources (e.g., large game, protected items such as shells and roots) through learning- and technology-intensive acquisition techniques; a large brain to enable massive amounts of learning and information storage; a long period of dependence before sexual maturity to support brain development and learning; a multigeneration system of resource transfers (flowing from grandparents to parents to children) to subsidize our slow, energetically costly developmental trajectory; provision of food, protection, and other resources from both mothers and fathers

(biparental investment), supported by marriage and long-term bonding between sexual partners; low mortality rates; and networks of cooperative arrangements among kin and unrelated individuals, which permit extensive food sharing and group coordination, further reducing mortality (Kaplan et al., 2007). Another major function of group cooperation throughout human evolution has been that of enabling collective aggression against other groups, from small-scale fights and raids to all-out wars between enemy coalitions (Bowles, 2009; Bowles & Gintis, 2011; Keeley, 1996; Tooby & Cosmides, 2010).

Social and Cultural Brains

The ever-increasing complexity of human social dynamics—including those generated by pair bonding, multigeneration transfers, extended cooperation, and conflict between and within groups—worked in synergy with foraging and tool-making to favor the evolution of larger brains and increased cognitive abilities (Alexander, 1989; Bailey & Geary, 2009; Dunbar, 1993; Dunbar & Shultz, 2007, 2010; Flinn et al., 2005). The social organization of foragers over the past 300,000 years was likely characterized by a nested hierarchy of groups, from small bands of 30–50 people to large tribes of perhaps 1,000 or 2,000 individuals, with an intermediate level (“clans” or “villages”) numbering 100–200 units. In short, the social world of our distant ancestors comprised hundreds of individuals, connected by various degrees of kinship and organized in groups and subgroups of variable size. This level of social complexity would not have been possible without language, which is possibly the most distinctive of all human traits (Birdsell, 1973; Dessalles, 2007; Dunbar, 1993; Pinker, 1994). Language and the cooperative sharing of information that it permits have entrenched humans in the *cognitive niche*—a unique foraging niche in which problem-solving based on cause–effect reasoning, transmitted knowledge, and social coordination are employed to overcome other organisms’ defenses in order to feed on them (examples are hunting large game, constructing weapons and traps, and detoxifying plants by cooking; Pinker, 2010; Tooby & DeVore, 1987).

The information-sharing skills that define the cognitive niche also enable cultural transmission on a massive scale (Dean et al., 2014; Richerson & Boyd, 2005; Tomasello, 2009). Humans possess a vast range of transmission

devices that include direct observation and imitation, teaching, storytelling, drawing, and (much more recently) reading and writing. Cultural transmission engenders a complex interplay between individual and social learning—and, more generally, between innovation and conformity, with their complementary costs and benefits for individuals and groups (Dean et al., 2014; Nettle, 2009a; Richerson & Boyd, 2005). The ability to build a cumulative culture has been a key factor in the propagation of human populations across the continents and their adaptation to a strikingly broad range of climatic and ecological conditions (Flinn et al., 2005; Nettle, 2009b).

Agriculture and Recent Human Evolution

In evolutionary terms, the invention of agriculture belongs to the very recent past. The earliest traces of farming in the Middle East go back to about 12,000 years, that is, roughly 500 generations ago. And yet, agriculture has been a major watershed in our evolutionary history. With the advent of farming, vital resources such as land and cattle could be accumulated, stolen, and defended; the outcome was a dramatic increase in social stratification, power/wealth inequality, and warfare, as well in the scale of cooperative networks. All these trends further intensified with the appearance of cities and states around 6,000 years ago (Kaplan et al., 2009; Powers & Lehmann, 2014). At the same time, food production increased by orders of magnitude, enabling unprecedented levels of demographic expansion and population density. The effects of these demographic changes reverberated through virtually every aspect of life—from diet and social interactions to the spread of infectious diseases (Cochran & Harpending, 2009).

Largely as a consequence of population growth, the rate of genetic change in humans seems to have accelerated dramatically over the past 10,000 years (Enard et al., 2014; Fu et al., 2013; Hawks et al., 2007). Genes showing evidence of recent evolutionary change include several that are expressed in the nervous system and some that have been associated with the risk for specific mental disorders (e.g., Grossman et al., 2013; López Herráez et al., 2009; Polimanti & Gelernter, 2017; Schrider & Kern, 2017). The once contentious notion that the evolution of human brain and behavior has accelerated in the recent past is becoming increasingly accepted. What is still controversial (despite being biologically plausible) is the idea that

populations inhabiting different ecologies and living in different social and cultural systems may have evolved systematic differences in genes that affect personality and cognition (Frost, 2011; Frost & Harpending, 2015; Wade, 2014; Winegard et al., 2017).

Using the Past to Understand the Present: Proximate and Ultimate Explanations

The central premise of evolutionary psychology is that our present minds embody the deep historical past, having been shaped across millions of years by the changing demands of the human ecology. Tinbergen (1963) famously pointed out that any biological system can be explained in four ways; that is, from the standpoint of *mechanism* (What is the system like? How does it work?); that of *ontogeny* or *development* (How does the system change over the organism's life course?); that of *phylogeny* (What is the system's history? How has it changed through evolution, and how does it differ between related species?); and, finally, that of *adaptation* (Why did the system evolve into its present form? What evolutionary advantages did it provide?).

Ontogenetic and mechanistic explanations concern the way an organism works in the present—in biological parlance, they are *proximate* explanations. *Ultimate* explanations (phylogeny and adaptation) consider the organism in relation to the past and the evolutionary forces that shaped its body and behavior. The two levels are pragmatically distinct but not mutually exclusive; on the contrary, proximate and ultimate questions are complementary and synergistic. Just as knowing the adaptive purpose of a mechanism can illuminate its functioning, understanding the mechanics and development of a trait constrains the range of plausible adaptive explanations for its evolution (Mayr, 1963; Scott-Phillips et al., 2011). The standard approach in psychology, medicine, and neuroscience is to focus almost exclusively on the proximate level of analysis. Evolutionary disciplines take the next logical step by asking questions about adaptation—which, as Darwin was the first to realize, ultimately reduce to questions about *reproduction*. The goal of the next section is to unpack the logic of natural selection and clarify the meaning of crucial terms such as fitness, adaptation, and maladaptation.

NATURAL SELECTION AND ADAPTATION

The concept of natural selection (Darwin, 1859) is the cornerstone of evolutionary biology. Natural selection occurs in a population of reproducing individuals whenever these conditions are met: (a) resources are limited, so that unconstrained reproduction is impossible; (b) individuals differ from one another in their morphological, physiological, and behavioral traits, or *phenotypes*; (c) at least some phenotypic traits correlate with an individual's ability to successfully reproduce, leaving more descendants in the next generation; and (d) phenotypes are inherited, meaning that they are transmitted to descendants with some reliability. When these conditions apply, individuals that are better able to reproduce leave more descendants; in turn, those descendants tend to carry the traits that favored reproduction in the preceding generations. Over time, successful traits tend to become more common—that is, they are selected for because of their positive effects on reproductive success or *fitness*. Traits that enhance fitness are called *adaptive*, while those that reduce it are called *maladaptive*; if a trait has no impact on fitness, it is considered *neutral* with respect to natural selection. Note that natural selection is an abstract process and does not require specific mechanisms of inheritance (such as DNA replication) in order to work. All it requires is the combination of heritable variation and differential reproduction based on that variation.

The basic measure of Darwinian fitness is an individual's lifetime reproductive success, or the total number of that individual's offspring that survive to maturity. (Fitness calculations become more complex when population growth and fluctuations in the environment are taken into account; see Hunt & Hodgson, 2010.) While surviving is usually a condition for reproduction, the catchphrase “survival of the fittest” is somewhat misleading, as survival without reproduction is an evolutionary dead end. It does not matter how well an organism is able to survive; if it fails to leave descendants, its enhanced survival abilities will not be passed down to the next generation and will eventually disappear. Organisms thus need to trade longer survival against increased reproduction, which is the real currency of evolution.

Inclusive Fitness

While individual reproduction is an important component of fitness, it still fails to fully capture the logic of natural selection. Indeed, many important and widespread phenomena—including altruism and self-sacrifice—involve a *loss* of individual fitness and may seem paradoxical or maladaptive on the face of it. Hamilton (1964) was the first to show that, in order to account for these apparent contradictions, the straightforward idea of fitness as individual reproduction must be replaced by the related but more sophisticated concept of *inclusive fitness*. Inclusive fitness theory—also known, somewhat improperly, as *kin selection*—provides a unified account of natural selection that includes social behavior in all its forms (Bourke, 2011; Grafen, 2009; West et al., 2007). In a nutshell, the theory shows that natural selection does not maximize individual reproductive success, but a different quantity called inclusive fitness. Inclusive fitness is the sum of the contribution of an individual (actor) to its own reproduction and its contribution to the reproduction of other individuals (recipients), the latter weighted by a coefficient of *relatedness* that summarizes the average genetic similarity between actor and recipient. In other words, an individual can increase its inclusive fitness in two ways: by increasing its own reproductive success (direct fitness) and by increasing that of genetically similar individuals (indirect fitness).

Relatedness is $r = 0$ between two individuals of a population picked at random and $r = 1$ between two genetic clones (such as identical twins). The relatedness coefficient may even become negative if the recipient is *less* genetically similar to the actor than a random member of the population (for example because the recipient belongs to a different population; see Grafen, 1985). Diploid organisms such as humans have two sets of chromosomes, one from each parent; as a result, relatedness between family members is high but considerably less than perfect. In simplified terms, the relatedness between parent and child is $r = 0.5$, the same as that between two full siblings; that between grandparent and grandchild is $r = 0.25$, the same as that between two half-siblings; and so on.

Inclusive fitness theory is encapsulated by the expression known as *Hamilton's rule*:

$$M + rN > 0$$

The rule states that a behavior—or any other phenotype with social effects—

will be selected for if its effect on the actor's fitness (M) and its relatedness-weighted effect on the recipient's fitness (rN) sum to a positive quantity. Note that M and N can be either positive (benefits) or negative (costs); likewise, r can range from positive to negative, as noted earlier. In the specific case of altruism, the effect of behavior is by definition a cost C for the actor and a benefit B for the recipient. By treating both B and C as positive quantities, one gets the standard version of Hamilton's rule for altruism:

$$rB > C$$

Altruism will be selected for if the benefit enjoyed by the recipient, weighted by the relatedness between actor and recipient (rB), is larger than the cost incurred by the actor (C). This means that costly altruistic behavior can evolve, provided that the benefit for the recipient and/or the relatedness between actor and recipient are sufficiently high. Also, biological altruism can only evolve when relatedness is positive ($r > 0$). However, the implications of inclusive fitness are not limited to altruism; Hamilton's rule can be applied to all kinds of social interactions, as shown in [Figure 1.1](#) (see Bourke, 2011; West et al., 2007, 2011).

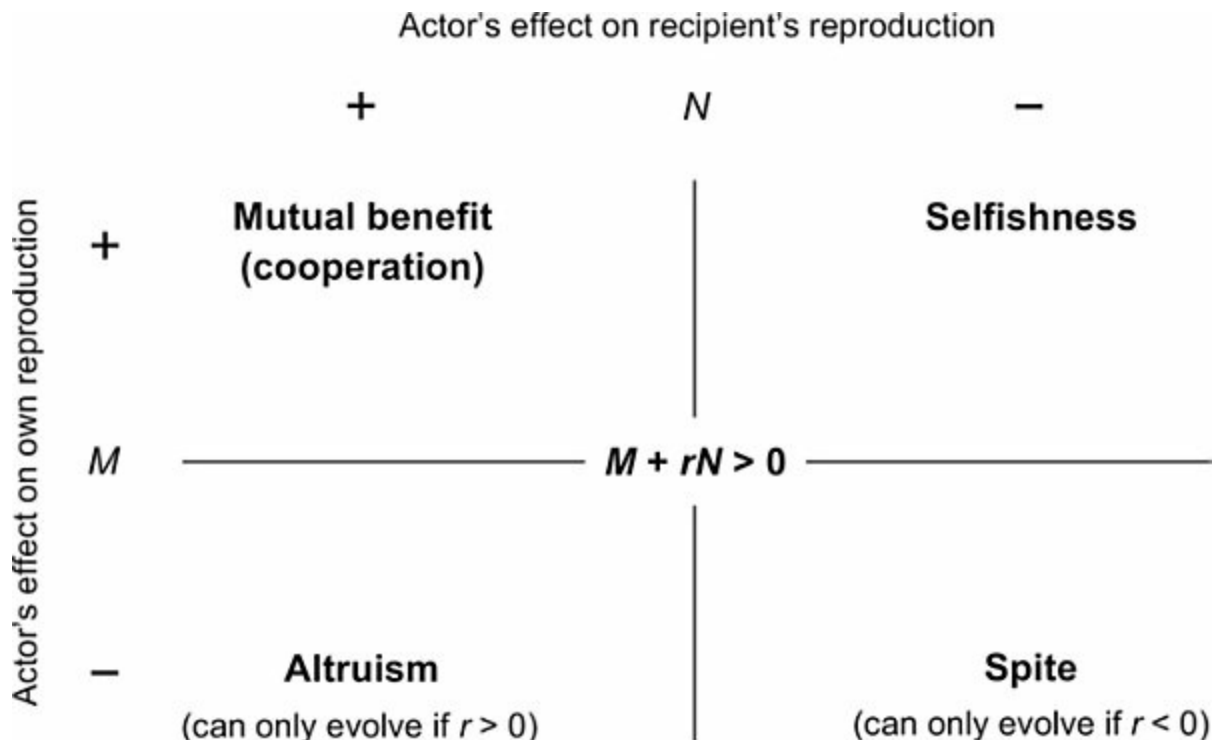


Figure 1.1. Hamilton's rule applied to the four main types of social interaction. A trait or behavior can be *mutually beneficial/cooperative* if both actor and recipient benefit from the interaction; *altruistic* if the actor incurs a cost in order to provide a benefit to the recipient; *selfish* if the actor benefits by inflicting a loss on the recipient; or *spiteful* if the actor incurs a cost in order to inflict a loss on the recipient. Whereas cooperation and selfishness can potentially evolve at any level of relatedness (depending on the exact balance of costs and benefits), altruism can only evolve if actor and recipient are positively related, and spite can only evolve if they are negatively related.

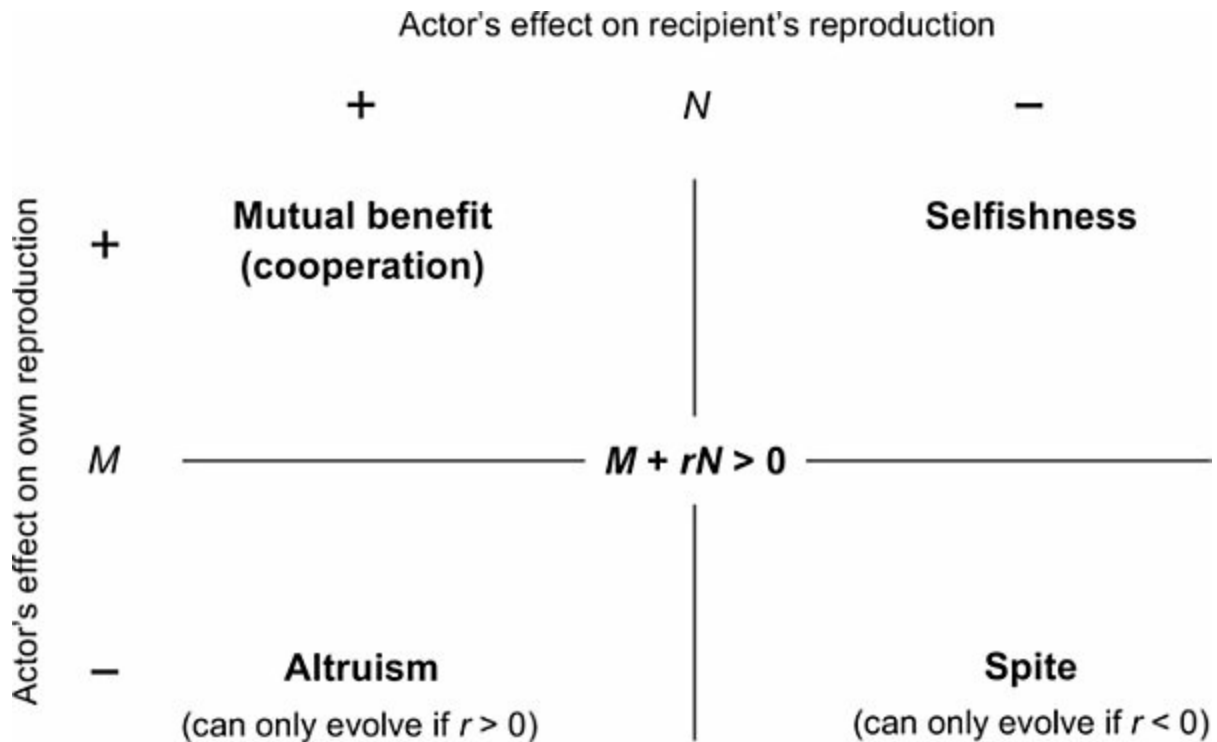


Figure 1.1. Hamilton's rule applied to the four main types of social interaction. A trait or behavior can be *mutually beneficial/cooperative* if both actor and recipient benefit from the interaction; *altruistic* if the actor incurs a cost in order to provide a benefit to the recipient; *selfish* if the actor benefits by inflicting a loss on the recipient; or *spiteful* if the actor incurs a cost in order to inflict a loss on the recipient. Whereas cooperation and selfishness can potentially evolve at any level of relatedness (depending on the exact balance of costs and benefits), altruism can only evolve if actor and recipient are positively related, and spite can only evolve if they are negatively related.

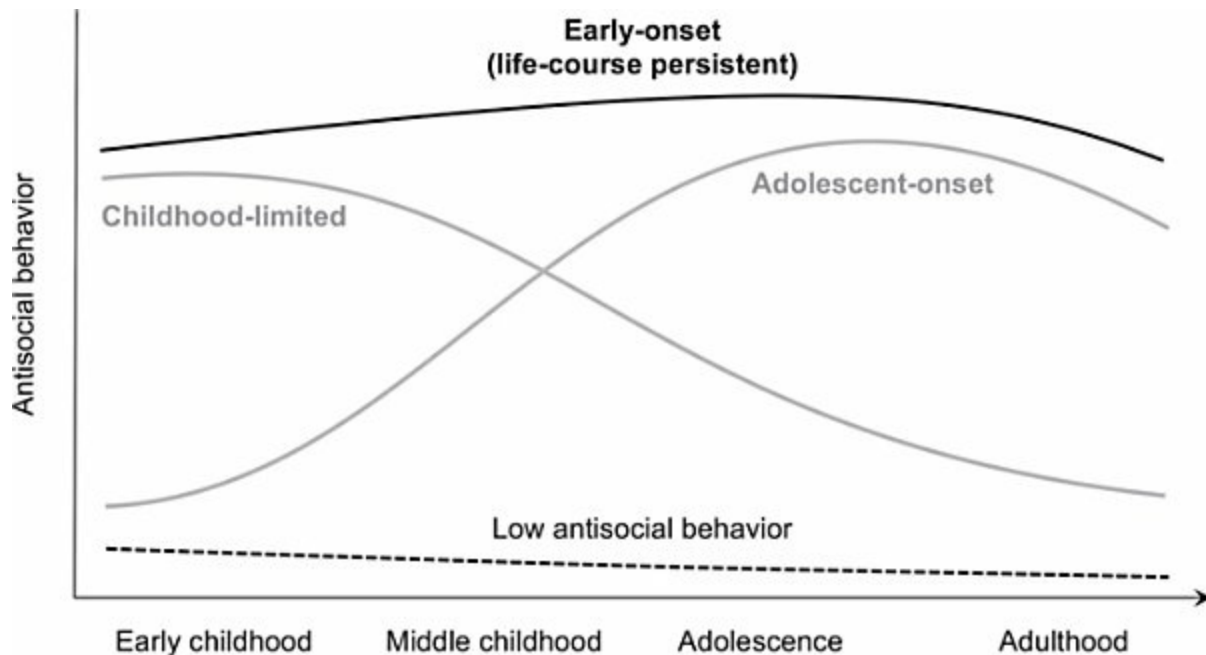


Figure 7.1. Common developmental trajectories of antisocial behavior. Note that “adolescent-onset” trajectories are characterized by a rapid increase of antisocial behavior during middle childhood.

Adaptive/Maladaptive Versus Desirable/Undesirable

In the evolutionary disciplines, the terms “adaptive” and “maladaptive” refer to the overall fitness effect of a trait. All traits have costs as well as benefits; to be adaptive, a trait does not have to be cost-free, but only to yield a positive net contribution to the organism’s inclusive fitness. These notions of adaptation and maladaptation contrast sharply with how the same terms are usually employed in psychology and psychiatry. In these disciplines, “adaptive” traits and behaviors are those that promote health, safety, subjective well-being, and mutually rewarding social relations; whereas socially undesirable, aversive, or health-damaging traits are viewed as “maladaptive” (see Nesse & Jackson, 2006). It is paramount to realize that these definitions of adaptation and maladaptation are conceptually orthogonal. Unsettling as it may be, the logic of natural selection promotes reproductive success rather than happiness, or even health per se (Cosmides & Tooby, 1999; Gluckman et al., 2011; Nesse, 2004a). Biologically adaptive traits may or may not be socially desirable or conducive to health and well-being; conversely, traits that consistently reduce well-being and adversely impact an individual’s health can be selected for—as long as they lead to

enhanced reproduction. In this book, I always refer to adaptive and maladaptive traits in the biological sense, and employ the terms “desirable” and “undesirable” to denote the implications of a trait for health, well-being, or social values.

Adaptations

By constantly weeding out unsuccessful variation, natural selection produces modifications of existing phenotypes, leading to the accumulation of characteristics that are organized to enhance survival and reproductive success. *Adaptations* are inherited and reliably developing phenotypes that have been selected for because of their causal role in enhancing the fitness of individuals that possess them (Williams, 1966). Through this process, adaptations acquire biological functions and the appearance of purposeful design: the immune system functions to protect organisms from pathogens, the heart functions as a pump for blood, and so on.

While adaptations are (by definition) a product of evolution, evolution does not always produce adaptations; likewise, not every characteristic of an organism is automatically an adaptation. For example, traits may become fixated in a population by *drift*, a random process by which neutral or even deleterious characteristics become more prevalent due to chance fluctuations in the frequency of traits. A neutral or weakly maladaptive trait may also spread by “hitchhiking” on another, positively selected trait to which it happens to be developmentally or genetically linked. In addition, many traits are not adaptations in themselves but rather *byproducts* of other adaptations. The sound the heart makes when it beats, the white color of bones, and the human chin are all nonfunctional byproducts of natural selection. A variety of approaches can be employed to identify adaptations and distinguish them from other types of traits. The most common methods include making phylogenetic comparisons between different species, including extinct ones (e.g., early hominids); estimating the fitness costs and benefits of a trait by measuring its associations with survival and reproduction; and building mathematical models of trait evolution that can be compared with empirical data or used to evaluate the plausibility of alternative hypothetical scenarios (Schmitt & Pilcher, 2004). Adaptations often show evidence of *special design*: the main features that indicate special design are economy, efficiency,

complexity, precision, specialization, and reliability in service of a trait's function. These features can be used to identify and “reverse-engineer” adaptations by framing them as solutions to specific biological problems (Tooby & Cosmides, 2015; Williams, 1966).

THE IMA PRINCIPLE

The logic of adaptation through natural selection has an important implication: as evolution proceeds, individual organisms are selected to develop and behave in ways that maximize their (expected) inclusive fitness. One does not have to assume that individuals are intentionally or consciously maximizing their fitness; they only need to function *as if* they were attempting to do so. This has been aptly described as the *individual-as-maximizing agent* or IMA principle (Grafen, 1999, 2006, 2009). Crucially, it can be shown that maximizing one's inclusive fitness is equivalent to maximizing the replication of one's genetic variants (*alleles*) in future generations, since—by definition—recipients who are more closely related are also more likely to carry copies of the actor's own alleles.

This deep equivalence is the source of Dawkin's much misunderstood metaphor of the “selfish gene”: individual organisms can be selected to behave altruistically (thus sacrificing their individual fitness), but only because doing so ultimately enhances the replication of their genes (Dawkins, 1976). Note that, in inclusive fitness theory, a gene is an abstract unit of inheritance, not a protein- or RNA-coding sequence of DNA as in molecular biology. For example, epigenetic mechanisms such as DNA methylation influence gene expression without modifying the underlying DNA sequence (Feil & Fraga, 2012; Jones, 2012; Ledón-Rettig et al., 2013). In some cases, epigenetic modifications can be transmitted from parent to offspring, even across multiple generations (Day & Bonduriansky, 2011; Ledón-Rettig et al., 2013; Meaney, 2010). Provided they follow the same pattern of transmission as the individual's DNA, stable epigenetic marks can be assimilated to genes from the standpoint of inclusive fitness and natural selection (Lu & Bourrat, 2017).

Of course, the IMA principle does not imply that a given individual will necessarily achieve high fitness; what is maximized by natural selection is only an individual's statistical *expectation*, not the realized outcome in any particular case. Even more crucially, optimization does not imply unconstrained “perfection” in the design of organisms and adaptations. Fitness maximization always takes place within the constraints and tradeoffs

imposed by the laws of physics, the characteristics of the environment, the phenotypic results of previous evolution, and the existence of evolutionary conflicts of interest (see later discussion). Furthermore, the IMA principle can only be expected to hold when the environment is sufficiently similar to that in which the organism evolved—which, in the case of contemporary humans, is often debatable (more on this in Chapter 5).

Group Selection

Some influential evolutionary scholars have invoked *group selection* as an alternative explanation of the evolution of cooperation and altruism in humans (e.g., Bowles & Gintis, 2011; Sober & Wilson, 1998; Wilson, 2012). The basic idea of modern group selection (or *multilevel selection*) is that competition for reproduction occurs simultaneously at two levels, between different groups within the broader population and between individuals within each group. While competition within groups tends to favor selfishness, competition between groups favors altruism; this means that altruism toward group members can be selected for as long as it is counterbalanced by a sufficient benefit for the group as a whole.

As it turns out, multilevel selection and inclusive fitness theory are mathematically equivalent and differ only in how they partition the costs and benefits of social traits (Marshall, 2011; West et al., 2007, 2011; West & Gardner, 2013). Whereas inclusive fitness partitions fitness effects between actors and recipients, multilevel selection partitions them between individuals and their broader social groups. The evolution of altruism by group selection can be explained just as well from the standpoint of inclusive fitness, since mechanisms of group formation increase relatedness within groups relative to that between groups. As a result, helping group members leads to an indirect fitness benefit that can be so strong as to override large costs in direct fitness. Depending on the question being asked, describing the evolution of a trait in terms of group selection rather than inclusive fitness may be mathematically more convenient or provide additional insight (Goodnight, 2013). The bottom line is that—despite frequent claims to the contrary—group/multilevel selection is not an alternative evolutionary process, but only a redescription of natural selection from a different point of view.

Evolutionary Conflicts

The very idea of evolution by natural selection is predicated on competition among alternative forms—including alternative alleles of the same gene that compete for replication and representation in future generations. Inclusive fitness theory shows how this fundamental conflict of interest can give rise to cooperation and even altruism given the right combination of factors (costs, benefits, and relatedness). At the same time, the logic of inclusive fitness shines a spotlight on the fact that cooperation is usually fragile and coexists with conflict in most kinds of social relationships (Bourke, 2011).

The clearest examples of evolutionary conflict are those involving predators and parasites, including the pathogens that constantly assault larger organisms. Here, relatedness is not an issue, and interactions between organisms are entirely driven by direct costs and benefits. Predator–prey and pathogen–host interactions usually evolve into endless “arms races” in which the evolution of better offensive mechanisms is followed by that of more effective defenses, which selects for even better offenses, and so on. However, other outcomes are possible as well. Pathogens may benefit by becoming *less* virulent if doing so helps them spread more effectively: this is especially likely when contagion occurs through personal contact (e.g., coughing, sexual intercourse) rather than through indirect channels such as insects or bodily waste. Sometimes, as in the case of gut microbes, pathogens turn into symbionts and engage in mutually beneficial exchanges with their hosts instead of harming them.

PARENT–OFFSPRING CONFLICT

At the opposite end of the spectrum, the highest levels of cooperation and altruism are found between close relatives. Humans are no exception to this pattern, as documented by the extensive cross-cultural evidence of favoritism toward kin (e.g., Burton-Cheller & Dunbar, 2014; Hooper et al., 2015; Madsen et al., 2007). Still, as long as relatedness is less than perfect ($r < 1$), Hamilton’s rule leaves room for conflicts of interest between family members. An especially pervasive kind of evolutionary conflict is *parent–offspring conflict* (Trivers, 1974; see Schlomer et al., 2011 for a nontechnical introduction). Leaving aside paternity uncertainty, the relatedness between a parent and its offspring has a constant value ($r = 0.5$ in diploid species). This implies that—all else being equal—parents are selected to value the welfare of all their offspring to the same degree. However, each offspring should

value *itself* ($r = 1$) more highly than its siblings and half-siblings ($r = 0.5$ or less). As a result, parents and offspring can be expected to disagree about the optimal distribution of resources among siblings: the parent would maximize its inclusive fitness by dividing resources (e.g., food) equally, but each sibling would maximize its inclusive fitness by getting a larger share of the pie at the expense of present and future siblings (Figure 1.2). The same logic applies to other types of parental investment, such as time and protection. In fact, *any* trait or behavior that benefits the offspring at a cost to the parent (or vice versa) can be subject to parent–offspring conflict.

In humans, parent–offspring conflict theory has been employed to explain mother–infant conflicts about the timing of weaning, conflicts between siblings in step-families, and even conflicts between young adults and their parents over the choice of romantic partners (e.g., Apostolou, 2014; Buunk et al., 2008; Fouts et al., 2005; Schlomer et al., 2010; van den Berg et al., 2013). Parent–offspring conflict on fetal nutrition and growth during pregnancy has a profound impact on prenatal physiology and is the ultimate explanation for the occurrence of gestational hypertension and diabetes (Haig, 1993; Schlomer et al., 2011). Of particular importance for psychopathology, the regulation of maternal cortisol and other stress-related hormones during pregnancy may also be subject to similar conflicts (Del Giudice, 2012a; Gangestad et al., 2012; Mokkonen et al., 2016).

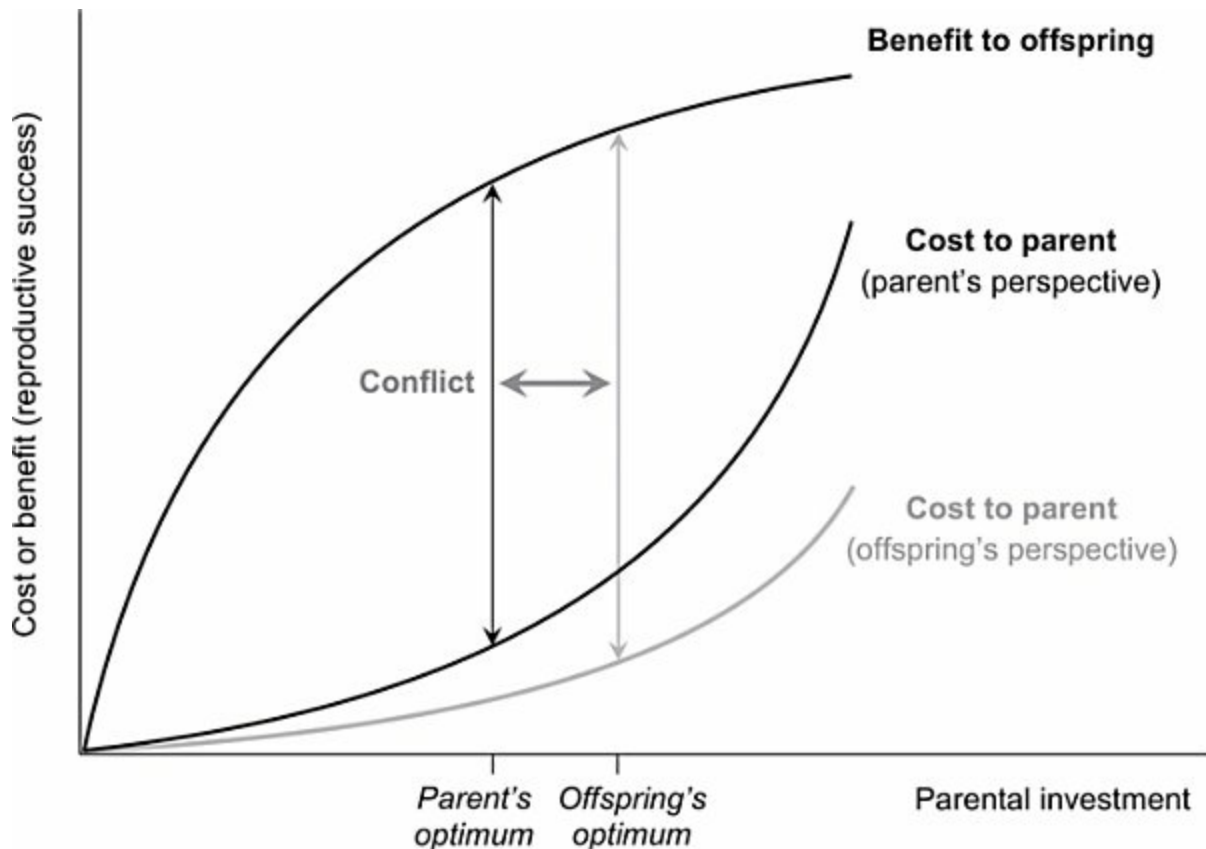


Figure 1.2. The logic of parent–offspring conflict. The offspring benefits from parental investment; at the same time, the parent incurs a reproductive cost because resources invested in the current offspring cannot be invested in future offspring. From the offspring’s perspective, this cost is discounted by the relatedness with future siblings. As a result, the optimal level of investment in the current offspring (i.e., the level that maximizes the difference between benefits and costs) is higher from the offspring’s perspective than from the parent’s perspective, engendering a conflict of interest between the two parties.

INTRAGENOMIC CONFLICT

Even within a single individual, different parts of the genome may not always share the same evolutionary interests, leading to *intragenomic conflict*. For instance, maternally derived genes (i.e., genes on the chromosomes inherited from the mother) have relatedness $r = 0.5$ with the maternally derived genes in each of the individual’s siblings. The situation is symmetric for paternal genes only if all siblings in a family can be expected to have the same father. With any degree of multiple paternity, some of the siblings are going to be half-siblings; from the perspective of paternally derived genes, this implies that relatedness between siblings is not 0.5 but a smaller quantity—much smaller if multiple paternity is common in a certain species or population. Because of this asymmetry in relatedness, maternally derived genes can

benefit by taking the mother's side in parent–offspring conflict (for example by reducing the growth rate of the fetus) while paternally derived genes benefit by siding with the offspring (Haig, 1997; Úbeda & Haig, 2003).

For most genes, the intrinsic conflict between patrilineal and matrilineal fitness is ultimately irrelevant because they carry no molecular memory of whether they were inherited from the mother or the father. However, at least a hundred genes in the human genome bear epigenetic marks that partially or completely inactivate them depending on their parent of origin (Barlow & Bartolomei, 2014; Davies et al., 2008; Gregg et al., 2010; Schalkwyk et al., 2010; Wilkins et al., 2016). These *imprinted genes* may evolve so as to shift the balance of parent–offspring conflict in favor of the mother (if maternally expressed) or the offspring (if paternally expressed; [Figure 1.3](#)). Many imprinted genes are expressed in the placenta, where they can participate in conflicts on parental investment—for instance by enhancing or suppressing fetal growth. Others are expressed in the brain, where they can potentially influence a wide range of behavioral traits subject to parent–offspring conflict (Haig, 2000; Úbeda & Gardner, 2010, 2011, 2012; Wilkins & Haig, 2003; for a review of alternative theories, see Patten et al., 2014). Intragenomic conflicts are not limited to imprinted genes but may also involve sexual chromosomes, mitochondrial genes, or “selfish” strands of DNA that—for various reasons—follow inheritance rules that differ from those of the rest of the genome (Burt & Trivers, 2006).

Because it breaks down the unity of the individual, intragenomic conflict works against the IMA principle: if different parts of the genome have divergent interests, selection cannot maximize any single measure of fitness. In most cases, however, the IMA principle remains valid as an approximation—because most genes within the genome have convergent interests being replicated and transmitted together, and because different sets of conflicting genes are likely to pull the phenotype in different directions, partly canceling out each other's effects (West & Gardner, 2013). Still, at a deeper level, the individual does not function as a fully self-consistent unit, but more like a compromise among competing interests within a collective of genetic “factions” (Haig, 2014).

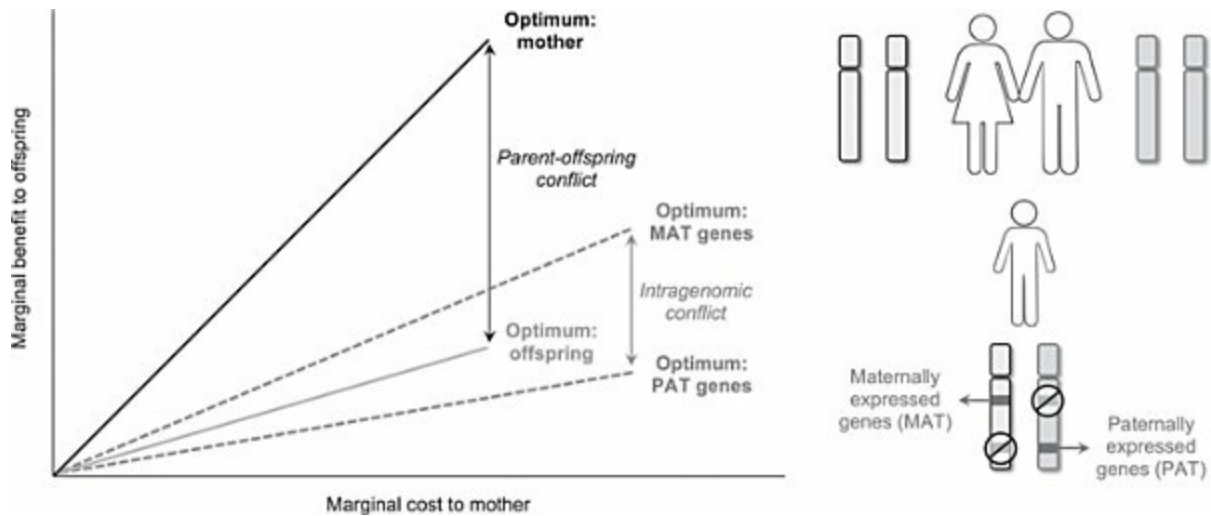


Figure 1.3. The logic of intragenomic conflict between imprinted genes. Because of parent–offspring conflict, the mother favors a higher benefit/cost ratio for parental investment than the offspring. However, paternally expressed genes within the offspring are less related to future siblings than maternally expressed genes; for this reason, they also favor a lower benefit/cost ratio than their maternally expressed counterparts. Paternally expressed genes may evolve as “resource enhancers” (genes that tend to maximize the investment provided by the mother, e.g., by increasing growth rate in fetuses or hunger in infants) whereas maternally expressed genes may evolve as “resources inhibitors” that counteract the action of resource enhancers. When both parents contribute to parental investment, more complex patterns may evolve; for example, if paternal investment increases over time, paternally expressed genes may behave as resource enhancers early in life but resource inhibitors later on (Úbeda, 2008).

SEX AND MATING

As in other sexual organisms, the biology of mating shapes human behavior in myriad ways. Social dynamics acquire much of their meaning from their direct and indirect connections with mating and reproduction; moreover, the selection forces that originate from sexual competition are ultimately responsible for many important differences between the psychologies of males and females.

Sexual Selection

In species that reproduce sexually, phenotypic traits can be selected for because they increase the quality or number of an individual’s mates (Darwin, 1871). This can happen in two ways: by making individuals

compete more effectively with rivals of the same sex, or by making them more attractive to potential mates of the opposite sex. Selection that arises from competition for mates is called *sexual selection* (Andersson & Simmons, 2006; Kuijper et al., 2012). While some authors draw a clear-cut distinction between natural selection (driven by survival) and sexual selection (driven by mating), it is more useful and logically consistent to define natural selection broadly in terms of inclusive fitness and treat sexual selection as a special case (Clutton-Brock & Huchard, 2013). Sexual selection is just one of many possible forms of *social selection*, in which traits are selected by other individuals through competition and choice (Lyon & Montgomerie, 2012; West-Eberhard, 1983). Examples of nonsexual social selection are the evolution of altruistic tendencies that increase one's attractiveness as a partner for cooperation and sharing, or the evolution of "cute" features that elicit interest in parents and make them more willing to protect their infants and children (Bogin, 1997; Locke & Bogin, 2006; Nesse, 2007).

Sexual selection drives the evolution of two main types of traits: *weapons* (traits used to physically compete with same-sex rivals) and *displays* (traits used to attract members of the opposite sex or repel same-sex rivals). In most species, males engage in physical competition and develop extravagant displays, while females play the role of the choosing sex. This is connected to the fact that females in those species provide most or all of the parental investment, including feeding, protection, and the physiological costs of eggs or pregnancy (Janicke et al., 2016; Kokko & Jennions, 2008; Trivers, 1972). However, several species exist in which males contribute to parental investment and mate choice is more reciprocal, and there are a minority of species in which sex roles are "reversed" (that is, females compete while males choose and provide for the offspring; Clutton-Brock, 2007; Janicke et al., 2016). In general, the sex that provides less parental care can reproduce at a higher rate and thus benefits more from investing time and energy in seeking multiple sexual partners. However, the evolutionary dynamics of sex roles are complex and depend on several interacting factors, such as uncertainty of paternity and the relative proportion of males to females in a population, or *sex ratio* (Kokko & Jennions, 2008). A common outcome of those dynamics is the evolution of alternative mating strategies, usually in males. For example, some males may fight for dominance or territory while others avoid competition and attempt to "sneak" copulations with unguarded females. Or, some may specialize in parenting while others become highly

competitive in the attempt to mate with as many partners as possible (Stiver & Alonzo, 2009; Taborsky & Brockmann, 2010).

The most fascinating products of sexual selection arise from mate choice. Selection based on sexual displays has a peculiar self-reinforcing nature: the displaying sex is constantly under selection to express more extreme versions of the trait, which in turn raises the bar for mate selection in the choosing sex. This process has the quality of an arms race between the displaying and choosing sex and explains the evolution of extravagant displays, such as the peacock's tail and the bowerbird's nest. The two basic mechanisms of selection via mate choice are *runaway* sexual selection (also known as Fisherian or "sexy sons" selection) and selection for *fitness indicators* (also known as "viability indicators", "genetic indicators", or "good genes" selection). In runaway selection, genes that determine the development of a trait in one sex become statistically linked with genes that control preference for the trait in the other sex, giving rise to a rapid and self-reinforcing selection dynamics. In fitness indicator selection, the trait becomes a display because it reveals the condition or "quality" of the bearer, including absence of deleterious genetic mutations (low *mutation load*), freedom from pathogens and other parasites (low *parasite load*), good nutritional status, and low exposure to stress during development. Note that the term "fitness" in this context indicates an individual's health and condition, not reproductive success or inclusive fitness. The various aspects of an individual's quality are often interdependent—for example, individuals who are healthier because of their low mutation load are often better able to secure high-quality food. Also, when both sexes contribute to parental investment, a mate in good condition is not only able to provide good genes for one's offspring but also high-quality investment—such as better food or better protection and defense of the young. Runaway and fitness indicator selection are best seen as extremes on a continuum, and even traits that originate from runaway processes tend to become fitness indicators as evolution proceeds (Chandler et al., 2013; Kokko et al., 2002; Kuijper et al., 2012).

A trait may function as a fitness indicator because it requires high levels of energy, strength, skill, health, or precise coordination in the expression of many different genes during development. Whatever the mechanism, a fitness indicator must be highly condition-dependent, which also means highly *vulnerable* to various sorts of disruptions—including deleterious mutations, infections, malnutrition, stress, and toxins (Geary, 2015). An indicator of

condition is only useful if it discriminates well between individuals of different quality (Figure 1.4); as a result, sexually selected traits tend to show an especially wide range of individual variation. When a fitness indicator is established, other traits may evolve as *amplifiers*. An amplifier is a trait that, while not a fitness indicator itself, functions to increase the condition-sensitivity of an indicator. For example, a decorative pattern on a bird's tail (the amplifier) may make it easier for partners and rivals to evaluate the length of the tail (the indicator; Castellano & Cermelli, 2010; Hasson, 1990).

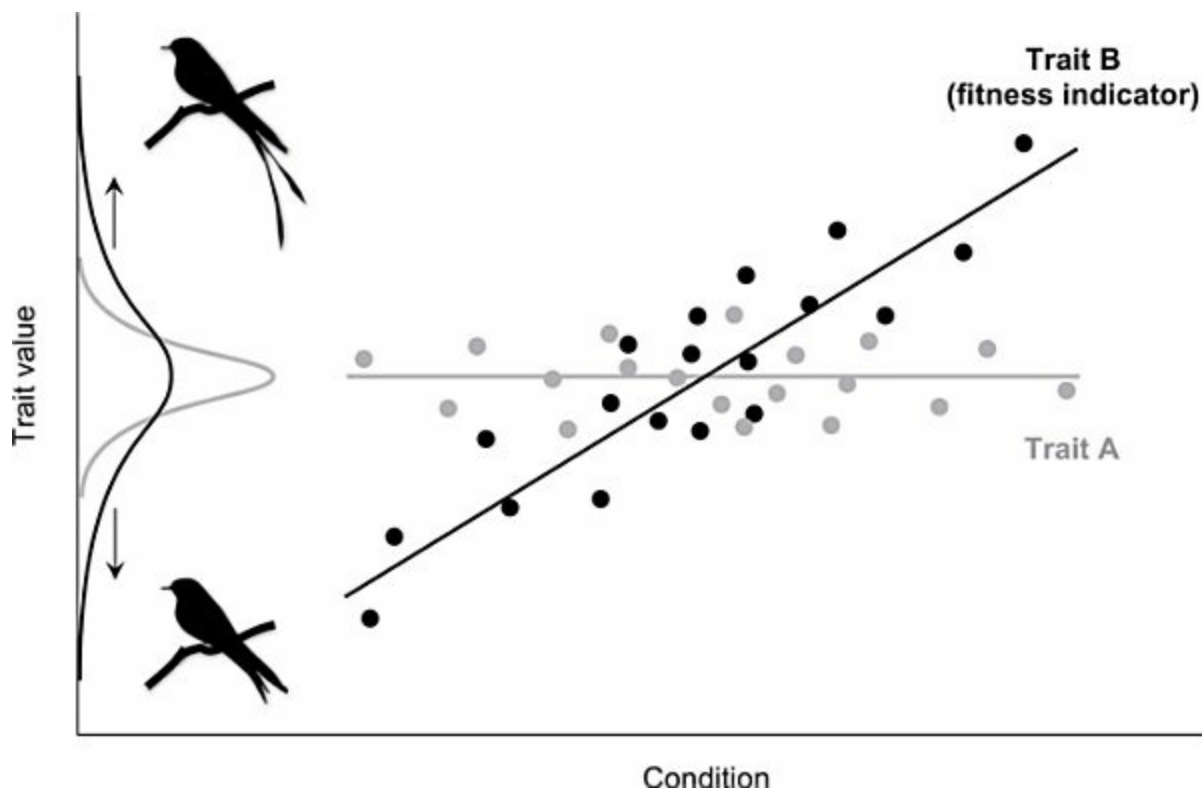


Figure 1.4. Fitness indicator traits are highly dependent on individual condition (e.g., low mutation and parasite load, good nutrition). As a result, they tend to show larger individual variation than other traits (see the distributions on the left).

OUTCOMES OF SEXUAL SELECTION

Sexual selection often drives the evolution of sex differences in morphology and behavior, but it can also produce similarities between the sexes, especially when partner choice is mutual rather than unilateral. Of course, the sexes may diverge even under mutual choice if males and females base their selection on different traits and qualities (Clutton-Brock, 2007). In the biological literature, discussions of sexual selection usually center on morphological traits (e.g., body size, weapons such as claws and horns,

ornaments) and on the behaviors involved in the utilization of those traits during contests and courtship (e.g., fighting rituals, courtship dances). However, *cognitive* traits are equally plausible targets of sexual selection. Selection for elaborate songs in bird translates into selection for learning abilities and memory capacity; in species with complex social behaviors, cognitive ability and social intelligence play important roles in both courtship and competition. Selection for cognitive abilities is especially likely to produce similarity between the sexes even when mate choice is unilateral, since the cognitive processes required to evaluate the quality of a display (e.g., a complex bird song) may have to be just as sophisticated as those needed to produce the display itself (Boogert et al., 2011; DeVoogd, 2004; Keagy et al., 2012; Miller, 2000).

Human Mating Systems

Human mating systems are remarkably flexible but, under the surface of cultural and ecological variation, there are some powerful common themes. To begin, *marriage* is a universal feature of human mating. Marriage in the traditional form is a social contract based on an enduring reproductive bond between a man and a woman. Cross-culturally, this contract implies a sexual division of labor: while both parents may contribute to provisioning, women disproportionately engage in childcare in every single society that has been studied (Geary, 2005, 2010; Walker et al., 2011). In all likelihood, pair bonding in our prehuman ancestors originally involved polygynous unions between a male and multiple females (Chapais, 2008, 2011). Today, about 80% of human societies allow polygyny; however, even in polygynous systems most marriages are monogamous, and only a minority of men actually have more than one wife. In contrast, polyandry (a woman marrying multiple men) is extremely rare and limited to marginal, resource-poor ecological contexts (Marlowe, 2003). The rate and intensity of polygyny increased dramatically after agriculture caused a rapid increase in social stratification and wealth inequality; the harems of ancient kings and emperors provide the most dramatic examples of this process (Kaplan et al., 2009; Walker et al., 2011). More recently, socially imposed monogamy has been replacing polygyny in a growing number of cultures, beginning in ancient Greece and Rome and spreading throughout Europe during the Middle Ages

(Henrich et al., 2012; MacDonald, 1995a; Scheidel, 2009). Across history and cultures, the reproductive success of men has been more variable than that of women, indicating a stronger intensity of sexual selection. Men are more likely than women to leave few or no descendants, but also to father really large numbers of offspring; this general pattern is attenuated in monogamous societies and in foragers, and amplified in polygynous and agrarian societies (Betzig, 2012; Brown et al., 2009).

Humans are clearly adapted for long-term pair bonding—whether monogamous or polygynous—with romantic love serving as a universal psychological mechanism of commitment. Consistent with a central role of pair bonding in offspring provisioning, marriages are most stable when men and women contribute about equally to family subsistence. At the same time, long-term bonding coexists with short-term, uncommitted sexual relationships and with the pervasive occurrence of extramarital affairs (Fletcher et al., 2015; Quinlan, 2008; Schmitt, 2015a). Paternity rates in humans are difficult to estimate, but the available evidence indicates high variability: extra-pair paternity may range from 1–2% (the most common scenario) to 10–20% depending on the specific cultural and social context (Anderson, 2006; Larmuseau et al., 2016; Scelza, 2011).

The interplay between long- and short-term mating is one of the defining features of human reproduction (Buss, 2003; Gangestad & Simpson, 2000; Schmitt, 2015a). Short-term mating with little or no parental investment can be highly adaptive for men, although there are also potential costs (including sexually transmitted diseases, retaliation from partners and relatives, and loss of reputation). Women benefit comparatively less from mating with many partners because their maximum rate of reproduction is severely limited by pregnancy and lactation. The main determinant of a woman's reproductive success is not her number of sexual partners, but the amount of resources—both material and social—that she can invest or leverage to the benefit of her children. However, short-term sexual relationships have advantages for women as well, from the direct exchange of sex for resources or protection to the chance of mating with a man of higher genetic quality than her long-term partner (sometimes called “gene shopping”). Also, short-term sexual encounters may facilitate mate switching for women who find themselves in unsatisfactory relationships (Gangestad & Simpson, 2000; Greiling & Buss, 2000).

PARENTAL CHOICE

Another peculiar aspect of human mating is the fact that parents and other close relatives often play an important role in the choice of marriage partners. Parental influence on mate choice can range from subtle to extreme, as in the case of arranged marriages. In both forager and agricultural societies, parental arrangement of marriage is the norm for daughters and a common occurrence for sons. While direct courtship is also widespread and may coexist with parental arrangement to some degree, it is often the case that parents have the last word on whether marriage proposals are accepted (Apostolou, 2010; Walker et al., 2011). Intriguingly, the preferences of parents and those of offspring can be expected to diverge somewhat as a result of parent–offspring conflict, promoting disagreement within families about the choice of the “right” partner (Apostolou, 2007, 2014; Buunk et al., 2008; van den Berg et al., 2013).

Mating Preferences and Strategies

Owing to the fact that both men and women contribute to parental investment, human mating has a strong component of mutual choice (Miller, 2000; Stewart-Williams & Thomas, 2013). When they look for long-term partners, both men and women value traits such as physical attractiveness (a reliable predictor of health; Kalick et al., 1998; Nedelec & Beaver, 2014; Tybur & Gangestad, 2011), intelligence, creativity, humor (a correlate of intelligence; Greengross & Miller, 2011; Hone et al., 2015), social status, industriousness, dependability, and kindness. At the same time, the sexes show some robust differences in their areas of competition and in their preferences for various desirable traits (Buss, 2003; Conroy-Beam et al., 2015; Davies & Shackelford, 2008).

On average, men show a stronger preference for attractiveness and beauty—especially for features that signal youth or high fecundity. (Following biological usage, in this book I use *fecundity* to denote an individual’s potential for reproduction and *fertility* to indicate that individual’s total number of offspring). In contrast, women are more attracted to indicators of status, resources (e.g., good financial prospects), and physical strength, and tend to prefer partners older than themselves (Gangestad & Scheyd, 2005; Grillot et al., 2014; Kenrick & Keefe, 1992; Shackelford et al., 2005). Preferences for creativity and humor production are also stronger in women

than in men (Beaussart et al., 2012; Clegg et al., 2011; Hone et al., 2015; Nettle & Clegg, 2006; Shackelford et al., 2005). Finally, women are more attracted to indicators of a man's willingness to invest in them and their children, such as kindness, generosity, parenting skills, and an interest in infants (Brase, 2006; La Cerra, 1994; Lukaszewski & Roney, 2010). The genetic evidence supports the hypothesis that sexual selection has been involved in the evolution of traits such as height, intelligence, and creativity (Verweij et al., 2014).

This general pattern of mating preferences is modulated by many personal, contextual, and social factors. For example, when men are looking for short-term sexual partners, they become much less selective about personality and tend to prioritize physical cues of current fecundity (e.g., breasts, buttocks) over cues of youth and health (e.g., facial features, symmetry; Confer et al., 2010; Currie & Little, 2009; Lu & Chang, 2012; Zelazniewicz & Pawlowski, 2011). Conversely, women prefer higher dominance and stronger masculine features in their short-term partners (Aitken et al., 2013; Little et al., 2011; Thornhill & Gangestad, 2008). In environments with high levels of pathogens, women's preference for indicators of health increases relative to that for dependability and paternal quality (Gangestad & Simpson, 2000; Lee & Zietsch, 2011; Tybur & Gangestad, 2011). Also, sex differences on several dimensions of mate preferences (e.g., social status, financial prospects, age) are attenuated in countries with higher gender equality in income and access to jobs (Zentner & Mitura, 2012; see also Schmitt, 2012). Depending on the social context, individual preferences may interact more or less strongly with those of parents. When parents evaluate their offspring's potential partners, they systematically place less importance on physical attractiveness and more importance on traits such as good family background, financial prospects, religiosity, and kindness—all indicators that prospective in-laws are going to invest in their offspring and bring benefits to the whole family (Apostolou, 2008a, 2008b, 2011, 2014; Buunk et al., 2008; Dubbs & Buunk, 2010; Perilloux et al., 2011).

ALTERNATIVE MATING STRATEGIES

Human mate preferences are complex and partially contradictory. As such, they leave room for a variety of viable mating strategies. This is especially true of male strategies, for two main reasons: first, female preferences are more complex and multidimensional; second, men have a wider range of pathways toward reproductive success, from long-term investment in a

woman and her children (“dad” strategies) to uncommitted relationships with multiple sexual partners (“cad” strategies). Throughout our evolutionary history, men’s strategies have changed and diversified with each major social transition. Considerable evidence indicates that dominance contests involving physical force and aggression have been the predominant mechanism of male sexual selection in ancestral humans and continue to play a significant role today (Puts, 2010; Hill, Hunt et al., 2013). This includes collective aggression between groups; in fact, capturing women was one of the main objectives of raids and wars in traditional societies (Kaplan et al., 2009). As paternal investment became an increasingly larger component of the human adaptive complex, some of the traits selected as weapons in male–male contests (e.g., upper body strength, throwing ability) may have been partially repurposed as provisioning tools in hunting, fishing, farming, and other forms of manual labor (see Geary, 2010).

The growing interdependence between group members caused by technological advancements and the consequent increase in role specialization—with multiple niches for hunters, tool-makers, healers, warriors, and many others—is likely to have had two important effects on mating competition. First, in a highly interdependent group, the benefits of dominance must be balanced against the need to maintain group cohesion and cooperation. By limiting the adaptive value of naked physical aggression, social interdependence must have intensified selection for social intelligence and political skills and probably favored nonphysical forms of competition (such as public speaking and “verbal duels”; Locke & Bogin, 2006; Puts, 2010). Second, technological and social interdependence make it possible to gain social status and resources by acquiring valuable skills, expertise, and knowledge, thus opening up alternative routes to mating success that do not depend on physical dominance.

In the evolutionary literature, this duality is often captured by the distinction between *dominance*, or social status acquired with the threat of force, and *prestige*, or social status freely conferred by others (Henrich & Gil-White, 2001). The importance of prestige in human societies likely increased with agriculture. As society became increasingly stratified, the number of specialized domains of skills and knowledge grew to unprecedented levels; also, for the first time, it was possible to accumulate wealth and pass it down to one’s descendants (e.g., Borgerhoff Mulder, 1987; Cronk, 1991). All the while, dominance competition escalated with the rise of chiefs and warlords

and the intensification of polygyny (Kaplan et al., 2009; Puts, 2010). More recently, socially imposed monogamy and state-managed law enforcement have contributed to reduce—but by no means eliminate—the benefits of dominance contests (Winegard et al., 2014). With the expansion of prestige-based competition during human evolution, the scope for direct mate choice by women (and their relatives) has increased considerably, likely resulting in stronger selection on display traits such as humor, verbal skills, and artistic creativity (Klasios, 2013; Miller, 2000; Puts, 2010). The partial divergence between parents' and offspring's preferences may also have contributed to select for *variation* in mating strategies: under parental choice, traits that are especially valuable to parents (such as religiosity and provisioning skills) can lead to reproductive success just as reliably as traits that appeal directly to the preferences of potential mates (Apostolou, 2014).

SOCIAL RELATIONSHIPS

Being an intensely social species, we spend most of our time interacting with other people in a variety of roles and situations. Human social dynamics are driven by the intricate interplay between cooperation and competition with other group members. From a biological standpoint, prosocial behaviors (such as cooperation, helping, and reciprocity) and aggressive, coercive behaviors can be seen as different types of strategies for gaining status and resources (Hawley, 2014a, 2014b). As described in the previous section, mating accounts for a large amount of competition revolving around dominance, prestige, and attractiveness. While direct mating competition in women is based more on physical attractiveness than on status and wealth, women do compete intensely with one another as they try to secure both material resources and social resources such as friends and allies (Benenson, 2013, 2014; Campbell, 2004; Geary et al., 2014).

Generally speaking, both males and females must balance “getting along” and “getting ahead” within their group; however, the two sexes solve this problem in somewhat different ways. Male groups are larger, show stable status hierarchies determined by one-on-one competition, and tend to be organized around shared goals and activities. Female groups are characterized by fluid, shifting hierarchies and high levels of emotional intimacy, sharing, and reciprocity. In female groups, overt competition and displays of

superiority are usually discouraged; in addition, female competition largely avoids physical confrontation and instead employs relational forms of aggression such as gossiping, dismissive nonverbal behaviors, and social exclusion or ostracism (Archer, 2009; Benenson, 2013, 2014; Campbell, 1999, 2004; Geary, 2010; Geary et al., 2003, 2014). While female aggression is usually tempered by caring and egalitarian concerns, there are many exceptions to this pattern; for example, wives in polygynous marriages often compete fiercely with one another for sexual attention, status within the family, and resources for their children (Geary et al., 2014).

Cooperation, Altruism, and Morality

Across societies, cooperation and altruism are greatly facilitated when they occur between close relatives (Burton-Cheller & Dunbar, 2014; Hooper et al., 2015; Madsen et al., 2007; Stewart-Williams, 2007). Even between nonrelatives, there is evidence that facial resemblance and other cues of potential relatedness lead to increased trust and altruistic dispositions (DeBruine et al., 2008; Lieberman et al., 2007); indeed, friends tend to be significantly more genetically similar to one another than to strangers (Christakis & Fowler, 2014). Of course, cooperation is not limited to relatives and often takes place between unrelated or weakly related individuals. People are highly sensitive to reciprocity in cooperative exchanges, and they are more likely to benefit someone if they expect that he or she will return the favor in time. This is known as *reciprocal altruism* (Trivers, 1971), although, from the standpoint of inclusive fitness, it is more correctly described as a kind of mutually beneficial cooperation (see Figure 1.1; West et al., 2007, 2011). Of course, reciprocity can be exploited by cheaters who enjoy its immediate benefits but then fail to reciprocate; this creates the opportunity for social and cognitive arms races in which increased ability to exploit and manipulate others coevolves with better skills for predicting behavior, detecting manipulation, and identifying and remembering cheaters (Cosmides & Tooby, 2015; Cosmides et al., 2010; Flinn et al., 2005).

LARGE-SCALE COOPERATION AND THE EVOLUTION OF MORAL NORMS

Human cooperation can take place on an extraordinarily large scale, easily involving hundreds, thousands, or even millions of individuals depending on the available technologies and means of communication. To achieve this feat,

we rely on a rich toolkit of psychological and cultural mechanisms. Two of these mechanisms are *reputation*—people constantly exchange information about the cooperative or selfish behavior of others—and *punishment* (West et al., 2011). Punishment occurs more frequently in larger communities, suggesting that it is an effective mechanism for promoting cooperation beyond the narrow circle of relatives and reciprocal exchange partners. The correlation with group size is especially strong for “altruistic” punishment, in which people willingly incur a personal cost to punish a violator (Henrich et al., 2010; Marlowe et al., 2008). On an even larger scale, *moral norms* seem to play a crucial role in suppressing—or at least regulating—selfish behavior and enabling large-scale cooperation with strangers. Moral reasoning is rooted in a multitude of evolved emotional mechanisms that regulate responses to harm (e.g., empathy), reactions to fairness and unfairness (e.g., gratitude, anger), hierarchical relations (e.g., respect, shame), and even contamination by pathogens and toxins (disgust; Haidt, 2007; Haidt & Kesebir, 2010; see Chapter 2). At the same time, human morality has been extended and organized with the cultural evolution of powerful institutions such as religions and systems of law (Norenzayan & Shariff, 2008; Norenzayan et al., 2016). Moral norms do more than just enable large-scale cooperation. First, they help maintain cohesion by marking the boundaries of the in-group. Second, they permit flexible social coordination by helping people choose sides in conflicts and disputes through moral judgment—that is, based on the moral rightness or wrongness of actions rather than simply on the identity of the conflicting parties. Potentially, this mechanism can prevent any particular individual from gaining too much power, while also avoiding the inevitable escalation of conflict that would result from the formation of rigid alliances (DeScioli & Kurzban, 2013; Haidt & Kesebir, 2010).

The biological and cultural evolution of moral norms creates more opportunities for social arms races. In this case, self-serving moral reasoning (e.g., hypocrisy, self-justification, advocating moral rules that promote one’s personal interest) can be expected to coevolve with a mentality of suspicion and with the ability to detect insincere or self-interested moral convictions (Haidt & Kesebir, 2010; Tooby & Cosmides, 2010). Also, reciprocity and moral norms provide the conditions for social selection on credible displays of altruism, fairness, empathy, and generosity. If individuals who display these traits are chosen more often as social partners, the resulting selection pressures can favor a rapid increase of altruistic and cooperative dispositions

within a population (Nesse, 2007). At the same time, selection for effective coordination between coalition members may explain pervasive group phenomena such as emotional contagion and conformity (Tooby & Cosmides, 2010). Social selection for altruism and other moral virtues is likely to overlap with sexual selection for the same traits. For example, men are more likely to engage in competitive displays of generosity, altruism, and self-sacrifice when they are being observed by women, consistent with an especially strong role for moral traits in female mate choice (Iredale et al., 2008; McAndrew & Perilloux, 2012; Miller, 2007; Tognetti et al., 2012).

Group Cohesion and Splitting

Group-splitting dynamics are a universal feature of human relationships, from forager bands to modern organizations such as business companies and political parties (Fox, 2004; Hart & Van Vugt, 2006; Walker & Hill, 2014). In group splitting or *fission*, some members leave the original group either to form a new group or to join an existing one. The trajectory of group splitting is accompanied by the escalation of conflict and by the progressive reinforcement of the subgroup's identity in opposition to that of the larger group. In traditional societies, common causes of splitting include stealing, free-riding, and other violations of reciprocity, as well as sexual affairs and—more generally—competition over scarce resources or mates. Competition between political or spiritual leaders often catalyzes group splitting: charismatic leaders play a dual role in this process as they foster cohesion within their faction while instigating conflict with adversaries (Bandy, 2004; Price & Stevens, 1999; Sani, 2008). Predictably, human groups often split at least partially along kinship lines, so that the members of a subgroup are more closely related with one another than with the rest of the group (Walker & Hill, 2014). Splitting becomes increasingly likely as groups become larger and the face-to-face mechanisms that promote cohesion and conflict resolution become progressively less effective. However, organized religions and other structured systems of moral norms can succeed in maintaining group cohesion on a much larger scale. Also, the prospect of conflict with a large competing coalition can powerfully motivate smaller groups to form alliances and merge with one another (Bandy, 2004; Haidt & Kesebir, 2010; Tooby & Cosmides, 2010).

DEVELOPMENT

Individuals change as they move through their life cycle—they grow, mature, decay, develop new traits and lose old ones. Some traits develop smoothly and gradually, while others change dramatically over a short period of time. The relation between evolution and development is deep and bidirectional: while all developmental processes are shaped by natural selection, selection can only modify the traits of an organism by acting on the developmental processes that build them (West-Eberhard, 2003). Any plausible model of human nature must take into account the developmental dimension of behavior and consider the interplay between physical and behavioral change over the life course.

Key Developmental Concepts

PLASTICITY AND CANALIZATION

As a rule, developmental processes are highly *canalized*—that is, they are designed to achieve their target phenotypic outcome despite fluctuations in the environment, perturbations of metabolic processes, mutations in an individual's genome or epigenome, and other sources of random noise. Canalization ensures that individuals of the same species will reliably develop the same basic anatomy and physiology despite being exposed to a wide range of external and internal conditions. At the same time, many phenotypes show some degree of plasticity and can vary in response to the specific environment encountered in development (Debat & David, 2001). While undirected plasticity is usually constrained by natural selection, organisms can evolve mechanisms of adaptive plasticity that respond to specific aspects of the environment in a way that maximizes fitness (DeWitt & Scheiner, 2004; Pigliucci, 2005; Schlichting & Pigliucci, 1998; see Chapter 3).

ONTOGENETIC AND DEFERRED ADAPTATIONS

A developmental perspective suggests a useful distinction between *ontogenetic* and *deferred* adaptations. Ontogenetic adaptations are selected to serve their adaptive function at a specific time in development and often disappear when they are no longer needed. Examples include the placenta, a fetal organ that provides nourishment and other vital functions during

prenatal development and is discarded immediately after birth; heat-producing brown fat in infants, which provides enhanced thermoregulation during a critical survival phase and is typically lost in adults; and infantile reflexes such as the suckling reflex (Bjorklund, 1997; Oppenheim, 1981). A subtler example is “neonatal imitation” of facial expressions in newborns. Neonatal imitation disappears around two months of age and—contrary to widespread belief—is not a precursor of later imitation skills. Instead, the primary function of this captivating phenomenon seems that of eliciting parents’ attention and facilitating early social interactions (Byrne, 2005). Many other aspects of cognition in infancy and childhood can be interpreted as ontogenetic adaptations with age-limited functions (Bjorklund & Ellis, 2014).

Deferred adaptations are phenotypes that appear in childhood but function—at least in part—to prepare children for adult behavior, and do not express their fitness-enhancing potential until sexual maturity (Hernández Blasi & Bjorklund, 2003). Play is an important example of deferred adaptation. A major function of play in humans and other mammals seems to be that of training the organism to deal with unexpected events, as well as regulating the long-term development of the neuromuscular system. Similarly, play parenting and play fighting are universal, sex-typical behaviors that foster the development of specialized adult skills in males and females (Byers & Walker, 1995; Geary, 2010; Pellegrini, 2013; Spinka et al., 2001).

Stages and Transitions

The human life cycle can be described as a sequence of functionally distinct stages, with transitions between stages usually controlled by hormonal mechanisms (Bogin & Smith, 1996). This pattern is especially clear before sexual maturity; by segmenting development into modular blocks, stages prevent interference between competing functions—for example by ensuring that body growth is complete before switching on the reproductive system—and permit tight coordination in the emergence of physical and behavioral adaptations. [Figure 1.5](#) summarizes the main stages of human development from conception to adulthood.

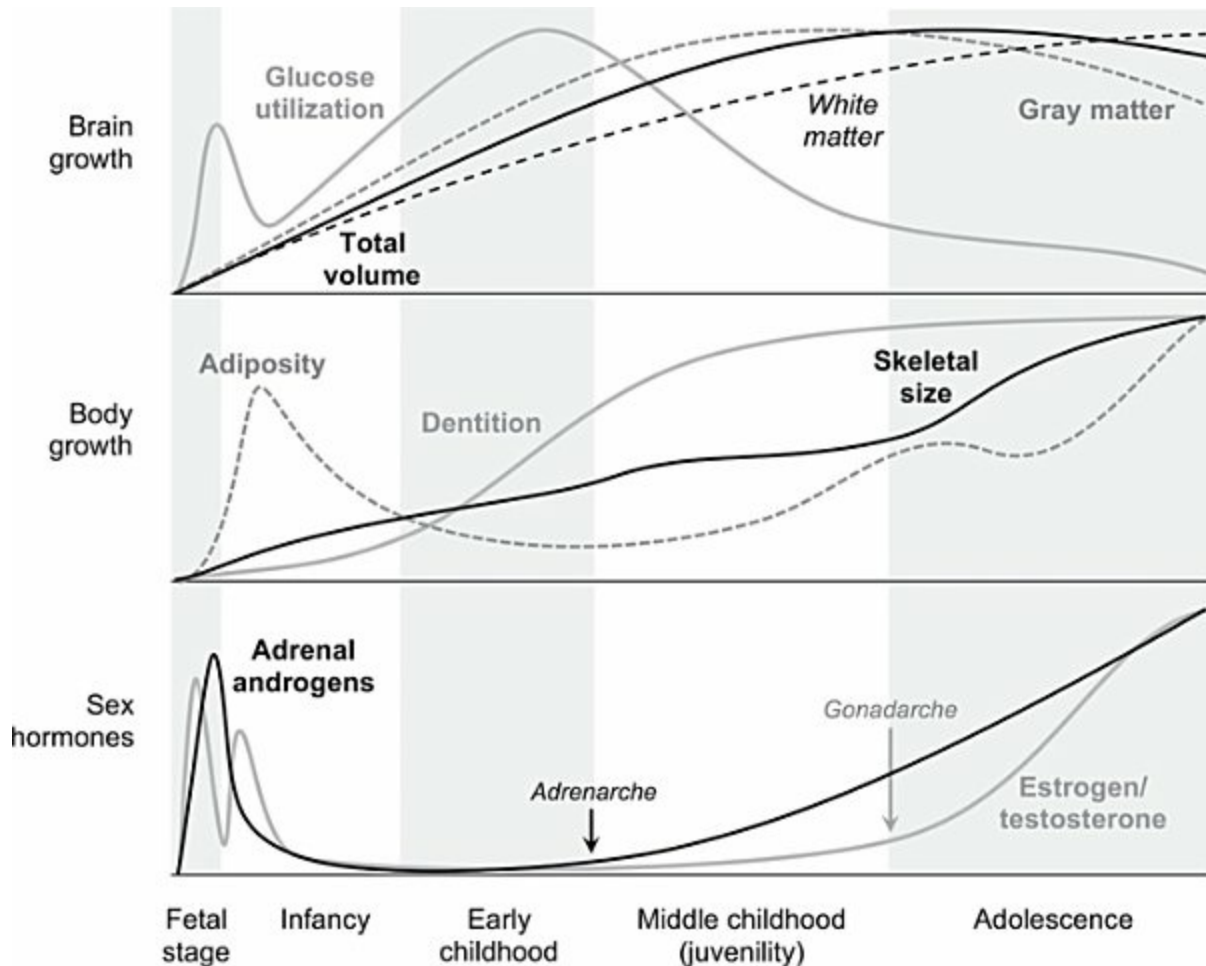


Figure 1.5. Developmental trajectories of human growth and sex hormones production, from conception to sexual maturity. Adapted with permission from Del Giudice (2014c).

FETAL DEVELOPMENT

The fetal stage is a critical phase of development; it takes place within the environment provided by the mother and further buffered by the placenta (which, as noted earlier, is made up of fetal tissue and can be regarded as an organ of the fetus). Fertilized eggs face extremely high mortality—between 30% and 70% of conceptions do not survive to birth—partly as a result of maternal “screening” mechanisms that trigger spontaneous abortion in the presence of genetically abnormal or malformed zygotes (Brosens et al., 2014; Bruckner et al., 2012; Forbes, 1997). Pregnancy involves a mixture of cooperation and conflict between the mother and the fetus, with a bidirectional network of physiological signals involved in the regulation of fetal nutrition, growth, and even maternal behavior (Del Giudice, 2012a; Haig, 1993, 2010; Schlomer et al., 2011; Sherman & Flaxman, 2002). The

initial development of the brain is a key function of the fetal stage. During pregnancy, sex hormones are produced both by the mother (mainly estrogen and small amounts of androgens) and by male fetuses during the second trimester (large amounts of androgens, particularly testosterone). Prenatal sex hormones have profound organizational effects on the developing brain and play an important role in the establishment of sex differences in behavior (Berenbaum & Beltz, 2011; Constantinescu & Hines, 2012; Hines, 2011).

INFANCY

In mammals, infancy is defined as the period in which the mother provides nourishment to her offspring through lactation. In traditional human societies, weaning occurs around 2–3 years of age. Lactation provides the resources for sustained brain growth and allows infants to accumulate fat stores that will support brain metabolism after weaning. Infancy is marked by rapid cognitive maturation, including the development of basic linguistic competence (Bogin & Smith, 1996; Locke & Bogin, 2006). Intriguingly, the first months of life are marked by a peak of testosterone production in male infants (Figure 1.5), which further contributes to sex-specific brain organization (Hines, 2011). After this early phase of activity, the testes become inactive until puberty and androgen levels drop accordingly.

Infants are fully dependent on their caregivers for survival; within the first year of life they form attachment relationships characterized by intense desire for closeness and separation anxiety, with the mother usually serving as the primary attachment figure (Cassidy, 2008; Marvin & Britner, 2008). And, indeed, infancy is a dangerous time: across cultures and history, 20–30% of newborns have died before the first year of life, with disease as the leading cause and another major contribution being infanticide and abandonment. Mortality remained high through childhood, so that 20–30% of the children who survived the first year still failed to reach puberty (Hrdy, 1999; Volk & Atkinson, 2013). Such high levels of mortality inevitably created strong selection pressures for ontogenetic adaptations designed to help fetuses, infants, and children survive the dangers of early development, even at a cost to later health or reproduction (Jones, 2009).

CHILDHOOD

Early childhood (about 3 to 7 years of age) is a uniquely human stage of development (Bogin & Smith, 1996). While other mammals begin to feed themselves right after weaning, humans spend several more years in which

they fully depend on adults for feeding. Childhood permits extended brain growth while freeing the mother from lactation, thus shortening the interval between pregnancies and increasing women's potential fertility. Glucose utilization by the brain peaks in early childhood, at around 4 years of age (Kuzawa et al., 2014). The transition from infancy to childhood is marked by a gradual increase in growth hormone, which then takes over as the main regulator of physical growth (Hochberg & Albertsson-Wikland, 2008). The cognitive breakthroughs of early childhood include the emergence of imitation and a dramatic increase in the ability to attribute mental states to other people, or *mentalizing* (also known as *folk psychology* and *theory of mind*; Bjorklund & Causey, 2018; Bjorklund & Ellis, 2014; see Chapter 2).

The next stage of human development is middle childhood (about 7 to 11 years in Western industrialized countries). Middle childhood is analogous to *juvility* in primates and other mammals—a stage in which the young is still sexually immature but no longer dependent on adults for feeding and protection (Bogin & Smith, 1996). Middle childhood starts with the eruption of the first permanent teeth and the awakening of the adrenal gland (*adrenarche*), which begins to secrete increasing amounts of adrenal androgens, such as dehydroepiandrosterone (DHEA) and its sulfate (DHEAS; Figure 1.5). Adrenal androgens have a broad range of effects; notably, they promote neural plasticity and shift the allocation of energy away from the brain and toward the accumulation of muscle and fat in preparation for puberty. They can also be converted to other androgens (e.g., testosterone) and estrogens in the brain, thus activating sexually differentiated pathways while further organizing brain development in a sex-specific manner (Campbell, 2006, 2011; Del Giudice et al., 2009; Del Giudice, 2014c; Kuzawa et al., 2014).

Middle childhood witnesses the emergence or intensification of sex differences in many domains—for example aggression and social play—as well as the first appearance of sexual and romantic attractions (Del Giudice, 2014c). The changes of middle childhood include dramatic increases in self-control and motor skills, which enables juveniles to help with domestic tasks such as foraging, preparing food, and taking care of younger siblings (Kramer, 2011). In total, middle childhood is a phase that combines intensive social learning and *integration* into one's group and culture with the emergence of social *competition* for status among peers (Del Giudice, 2014c). It also marks the first major peak in the onset of psychopathology

(Costello et al., 2003; Kessler et al., 2005).

ADOLESCENCE

Adolescence begins with a characteristic growth spurt and the onset of sex hormone production by the gonads (*gonadarche*). This is a phase of rapid physical and behavioral transformation that ends with the completion of growth and the achievement of full sexual maturity. In industrialized countries, the mean age of the first period in girls (*menarche*) has decreased from 15–17 to 12–13 years owing to improved nutrition and health (Parent et al., 2003). Girls remain infecund for some years after menarche, whereas boys become fecund shortly after the peak of their growth spurt. As in middle childhood, rising levels of sex hormones during puberty seem to have both activational and organizational roles on brain function (Bogin & Smith, 1996; Hines, 2011; Locke & Bogin, 2006; Romeo, 2003). Sexual motivation and status competition become central in adolescence; this motivational shift is paralleled by enhanced sensitivity to social cues and evaluation by peers, a dramatic rise in risk-taking behavior (especially in boys), and a surge in parent–child conflicts (Blakemore & Robbins, 2012; Ellis et al., 2012; Steinberg et al., 2008). In adolescence, the brain as a whole becomes smaller due to the loss of gray matter that accompanies cortical maturation; however, the development of axonal connections (reflected in increasing white matter volume) continues well into adulthood (Figure 1.5; Blakemore et al., 2010; Casey et al., 2011).

ADULTHOOD AND SENESCENCE

While the segmentation of the life cycle becomes less clear-cut after sexual maturity, several important transitions take place in adulthood. For example, the birth of a child is associated with extensive hormonal and behavioral changes in both parents (Edelstein et al., 2015; Gettler, 2014). From this vantage point, parenthood can be viewed as a functionally meaningful life stage even if it is not tied to a particular age. Another key transition for women is that of menopause, which marks the end of their reproductive window. Some aspects of behavior in old age may have been selected for in the context of specific human traits such as multigenerational provision and helping by grandparents (Gurven & Kaplan, 2009).

After maturity, bodily and brain functions gradually deteriorate as part of *senescence*, resulting in a pattern of increasing mortality over time (Jones et al., 2014). Even if senescence impairs health and accelerates death,

evolutionary models show that it can be optimal for organisms to allow their body to decay if later mortality is compensated for by increased reproduction earlier in life. In other words, senescence is not something that organisms endure passively, but (at least in part) an evolved design feature that depends on a specific balance of fitness costs and benefits (Kaplan & Robson, 2009; Kirkwood, 1990). The evolution of senescence powerfully illustrates the distinction between adaptive and desirable traits and underscores the role of natural selection in shaping the whole life cycle, from conception to death.

The Evolved Mind

In modern psychological terms, the mind can be described as a network of information-processing mechanisms whose function is to control and regulate behavior. These mechanisms are mainly instantiated in the brain, although some processing may take place in the rest of the nervous system and even outside of it (e.g., the immune system collects and integrates information about pathogen threats and can modulate brain functioning via hormonal signals). The mechanisms that enable and govern our behavior today have been shaped by the ecology and behavior of our ancestors across countless generations; the mind/brain can then be studied as an evolved “computational organ”—or, more precisely, a collection of specialized organs that perform various kinds of computations (Barrett, 2015; Carruthers, 2006). In this chapter, I introduce the idea that psychological mechanisms are adaptations and consider some of its most important implications. I then examine three aspects of mental functioning that are especially relevant to the disorders discussed later in the book: motivation and emotion, decision-making and self-regulation, and the distinction between mentalistic and mechanistic cognition. Throughout the chapter, I focus on the biological and computational functions of psychological mechanisms and touch only briefly on their localization in the brain. I conclude with an overview of key neurobiological systems, from brain monoamines such as dopamine and serotonin to sex- and stress-related hormones.

PSYCHOLOGICAL MECHANISMS AS ADAPTATIONS

As discussed in Chapter 1, the individual-as-maximizing agent (IMA) principle predicts that an organism’s behavior will tend to maximize its expected fitness, at least in environments sufficiently similar to that in which the behavior originally evolved. However, fitness cannot be maximized

directly but only indirectly by being successful at foraging, avoiding pathogens and predators, finding and attracting mates, and so on. Moreover, inclusive fitness does not map neatly on individual reproductive success, and some of its components (e.g., indirect or subtle effects on the reproduction of related recipients) may be difficult or impossible to observe within a lifetime. As a result, any individual trying to directly maximize its own fitness would face a hopeless task. The black-box observation that behavior tends to maximize fitness does not imply that organisms possess a general “fitness maximization mechanism”; rather, behavior is likely to be controlled by multiple systems that specialize in reaching narrower fitness-related goals such as finding and choosing food or improving one’s social status. A key tenet of evolutionary psychology is that people should be viewed as “adaptation executors” rather than direct “fitness maximizers” (Tooby & Cosmides, 1992, 2015).

Another important reason to focus on psychological adaptations is that they are the best candidates for the truly universal features of human nature. While human behavior shows many regularities, specific behavioral patterns are usually variable and contingent on learning, ecological and social conditions, and individual differences. Psychological mechanisms embody the computational rules that enable learning and behavioral flexibility; thus, the basic design features of a mechanism—or, more precisely, the evolved processes that control and regulate its development—can be universal even if the resulting behaviors are highly variable and context-dependent (Barrett, 2015; Bjorklund et al., 2007; Tooby & Cosmides, 1992).

The Environment of Evolutionary Adaptedness

Biological adaptation is a historical concept; in essence, all claims about adaptation are claims about the past. Traits that promote fitness in a certain environment may become neutral or maladaptive if the environment changes. While natural selection is expected—all else being equal—to weed out traits that have become detrimental to fitness, the process may often take a long time. This generates the potential for *mismatch* between an organism’s adaptations and its present environment.

The environment in which a given trait evolved is often referred to as its *environment of evolutionary adaptedness* (EEA). Despite some

misrepresentations in the literature, the EEA is not a specific time or place (e.g., the African savannah or the Pleistocene), but a statistical composite of those aspects of the environment that have shaped the evolution of a given trait across space, time, and generations. For this reason, different adaptations within the same organism can have widely different EEAs (Durrant & Ellis, 2003; Tooby & Cosmides, 1990). Consider two human adaptations: language and shivering in response to cold. While language originated in the human lineage no more than 2 million years ago, the shivering response has a much more ancient history and is shared with other warm-blooded vertebrates. Likewise, the dimensions of the environment that shaped shivering and language show virtually no overlap—factors such as external temperature and metabolic rates are crucial for the evolution of shivering but irrelevant to that of language. Because the EEA is not a specific environment, it does not have to be static; on the contrary, the amount and type of environmental variation experienced during the evolution of a trait is an important component of that trait’s EEA. For example, metabolic processes can evolve to maximize survival in unpredictable environments in which food abundance is suddenly followed by starvation.

Functional Specialization

In the most general sense, the overarching function of the mind/brain as a whole is to generate predictions in the service of action (Bar, 2007; Barrett, 2015; Clark, 2013). At the same time, there are strong reasons to believe that the mind—including “higher” mental functions such as reasoning and decision-making—is composed of a large number of evolved computational mechanisms, each specialized to process a specific kind of input (Barrett, 2015; Carruthers, 2006). Note that, in the broad sense employed here, computation refers to the manipulation of information regardless of its format (e.g., symbolic representations, distributed activations within a neural network; see Piccinini & Scarantino, 2011). Functional specialization (often discussed as *modularity* or *domain specificity*) is a common outcome of natural selection because it promotes computational efficiency, increases robustness (the mind as a whole can function even if a particular mechanism is damaged or perturbed), and facilitates rapid evolution as it allows selection to modify and fine tune the design of a mechanism without affecting the rest

of the system (Tooby & Cosmides, 1990, 1992, 2015). As a rule, psychological mechanisms do not work as indivisible units but can be further decomposed into a number of simpler elements, some of which may be shared with other mechanisms. Because of the nested and partially overlapping nature of mental adaptations, psychological taxonomies are necessarily “fuzzy,” and there are multiple valid ways of drawing boundaries between different systems and subsystems.

The question of how psychological mechanisms interact to produce flexible, adaptive behavior is still largely unanswered. In principle, the components of mind might be arranged in a hierarchical structure whereby higher level mechanisms monitor and control the activity of lower level ones, or in a horizontal structure whereby multiple interacting mechanisms work in parallel and compete for the control of behavior (e.g., Minsky, 1988; see Barrett, 2015, Carruthers, 2006). Still another alternative is a “soft” hierarchy in which interactions are mainly horizontal, but some mechanisms exert asymmetric influences or even “veto power” over others (e.g., through unidirectional inhibitory inputs).

THE NATURE OF SPECIALIZED MECHANISMS

Over the years, the issue of specialization has generated much debate but also a lot of conceptual confusion. Contrary to widespread misconceptions, functional specialization does *not* entail that a mechanism will process its input in an automatic or reflex-like fashion; that it will be unresponsive to contextual factors and conscious control; that it will be “encapsulated” or impermeable to information exchange with other mechanisms; that it will be present and functional from birth; or that it will follow innate rules that cannot be modified by learning. Instead, the operating characteristics of any given mechanisms should depend on its particular evolved function; for instance, specialized decision-making mechanisms may be *designed* to integrate a broad range of inputs, interact with other systems, and respond in a flexible way to contextual factors. Other mechanisms may respond automatically and inflexibly but depend heavily on learning—a dramatic example of the latter is Pavlovian (classical) aversive conditioning. More generally, learning is by no means antithetical to evolved specialization; on the contrary, psychological mechanisms may require a considerable amount of experience and environmental input to fully acquire their functionality. Conversely, adaptive learning processes must embody at least some assumptions about the expected structure of the world and/or be biased to

attend to certain kinds of stimuli—in other words, they must be functionally specialized to some degree (see Barrett, 2015; Barrett & Kurzban, 2006).

The popularity of the idea that specialized mechanisms are rapid, automatic, nonconscious, and associative whereas domain-general mechanisms are slow, controlled, conscious, and rule-based can be linked to the rise of dual-systems theories of cognition. In their early versions, these theories contrasted a rapid and intuitive “System 1” with a slow and deliberative “System 2” (Evans, 2008). More recently, however, dual-systems theorists have considerably softened the distinction and acknowledged that psychological mechanisms may show combinations of both types of attributes (e.g., both automatic *and* rule-based). Accordingly, the classic System 1–System 2 distinction has been recast as a much narrower contrast between multiple “Type 1” systems that require little or no access to working memory and multiple “Type 2” systems that rely on working memory to support hypothetical thinking and the simulation of mental scenarios (Evans & Stanovich, 2013).

In current evolutionary thinking, the domain of a mechanism is defined by the formal properties of its input and does not have to correspond to a narrow type of *content* such as food items, predators, or sexual partners. Psychological mechanisms can specialize to process abstract properties such as the temporal or logical sequence of a chain of events; also, their domain of operation can be extremely wide (e.g., “all physical objects” or “all rewarding stimuli”). Thus, a given mechanism can serve multiple functions and be recruited to perform different kinds of tasks (Anderson, 2010). On this view, the distinction between domain-specific and domain-general mechanisms loses meaning: *all* evolved mechanisms have a domain, however wide and abstractly defined (Barrett & Kurzban, 2006). This meaning of functional specialization is broader than the one usually found in neuroscience, where the term is employed to describe brain areas that specifically respond to certain content categories (e.g., faces, places, tools) or certain types of task (e.g., language comprehension, spatial orientation; Kanwisher, 2010). It is also useful to distinguish between a mechanism’s *proper* domain—the input it was selected to process in the EEA—and its *actual* domain: that is, the input it processes in the present (Sperber, 1994). Sensory mechanisms that evolved to detect high-calorie sugars in food can now be activated by artificial sweeteners without any nutritive value; mentalizing processes that evolved in the context of face-to-face relationships can be applied to fictional characters in novels and movies. Over time, the proper domain of a

mechanism and the range of functions it performs can change as individuals are exposed to novel selection pressures.

It is important to stress that functional modularity is not the same as *anatomical* modularity at the macroscopic level of visible brain structures. In neuroscience, modules are usually defined as clusters of brain regions that show strong anatomical connections, correlated variation in size, or correlated activity patterns (e.g., Gómez-Robles et al., 2014; Meunier et al., 2010). Since neural communication is more efficient at shorter distances, it is reasonable to expect some overlap between functional and anatomical modularity. To be sure, the human brain has a markedly modular organization, which seems to have facilitated the semi-independent evolution of different regions and structures (Barger et al., 2014; Gómez-Robles et al., 2014; Kanwisher, 2010; Wig, 2017). However, there is no reason that psychological mechanisms should always be localized within narrow areas of the brain; on the contrary, some mechanisms may be widely distributed across brain regions and include both cortical and subcortical structures. The analogy with bodily organs is a useful one. Whereas the liver is highly localized, the immune system is spread across the whole body; the circulatory system is similarly distributed, but it also includes some localized structures (e.g., the heart). In all these examples, a mechanism's degree of localization is ultimately dictated by its biological function.

The flip side of specialization is that different psychological mechanisms may partly rely on the same processes. Powerful computational principles such as Bayesian inference (see later discussion) and reinforcement learning can be successfully applied to a wide range of information-processing tasks including perception, reasoning, decision-making, and action control (Clark, 2013; Dayan, 2012; Dayan & Daw, 2008; Geisler & Diehl, 2003; Gopnik & Bonawitz, 2015; Kacelink, 2012; Vilares & Körding, 2011). In this sense, those computational principles can be regarded as domain-general, even if the mechanisms that employ them have their own specialized domains (Carruthers, 2006). Also, most models of the mind postulate the existence of shared cognitive workspaces (e.g., working memory) that can be accessed by multiple mechanisms to carry out their specific tasks.

MOTIVATION AND EMOTION

Motivational systems (also called *behavioral systems* in the ethological literature) are psychological mechanisms that organize behavior by orienting an individual toward biologically significant goals, tracking its progress toward those goals, responding to opportunities or obstacles, and promoting learning of goal-related strategies and skills (see Aunger & Curtis, 2013; Carruthers, 2006). In some cases, the goal of a motivational system is to keep a certain parameter close to its set point (homeostasis), as in the regulation of drinking and fluid intake. An example of social homeostasis comes from infant attachment, where a key goal of the system is to maintain physical proximity to the caregiver. In other instances, motivational systems are designed to flexibly adjust their goals in response to circumstances, as in the regulation of mating and dominance behaviors (e.g., the adaptive goals of high-ranking individuals are often different from those of low-ranking ones).

Importantly, the goals of a system need not be consciously represented by the individual. In extreme case, they may not be represented at all; bacteria manage to effectively swim toward nutrients without any internal representations of the reason for their behavior—or even of the direction in which they are swimming (see Dennett, 2009). In more complex organisms, motivational systems may rely on multiple sources of information about the state of the individual and its environment. For example, the pituitary gland constantly collects and integrates information about the volume and dilution of the blood, which in turn informs the system that controls thirst and fluid intake. The same logic applies to social motivations. In Chapter 1, I discussed how coefficients of relatedness play a crucial role in the cost-benefit analysis of cooperation and altruism. In the real world, relatedness is not a known quantity, but has to be estimated by combining various sources of information about other individuals. Some relatedness cues are weak and probabilistic (e.g., physical resemblance), while others are clear-cut and very accurate (e.g., having been nursed by the same mother). In humans, there is evidence that relatedness estimates are computed based on multiple cues and used to modulate altruistic behavior, sexual attraction, and other social motivations (Lieberman et al., 2007; Tooby et al., 2008). The information summaries employed by motivational systems can be described as *internal regulatory variables* (Tooby & Cosmides, 2008, 2015; Tooby et al., 2008). In all likelihood, most of the regulatory variables in the human brain are computed outside of consciousness and are largely or fully inaccessible to introspection.

While organisms have multiple biological goals, they cannot pursue all of them at the same time. Determining and managing priorities is thus a central task of motivational mechanisms. A key component of any motivational system is the set of rules that determine when and how strongly it will become activated, based on internal and/or external cues and the associated regulatory variables. A system's "rules of engagement" may be more or less complex and may depend on the activation state of other systems. The motivation to eat is partly determined by an individual's internal state (e.g., stomach fullness, blood glucose), but the sight and smell of delicious food may induce hunger even when internal cues alone would not. Conversely, hunger may be suppressed when other high-priority systems are activated, for example in presence of an immediate danger such as a venomous snake. Finally, motivational systems receive input from cognitive appraisals—a snake in a glass box may evoke a very different response than the same snake crawling on the floor. When two or more systems with incompatible goals become activated at the same time, the resulting motivational conflict may require the intervention of specialized regulatory mechanisms, as I discuss later in this chapter.

Emotions

When motivational systems are activated, they trigger the onset of emotions (such as fear and joy) and other feelings such as hunger, sexual arousal, and safety. The line between emotions and feelings is a blurred one and may have to do more with the conventions of language than with biological and psychological reality. Both emotions and feelings have affective *valence*—they are either subjectively positive (rewarding) or negative (aversive)—and both contribute to regulate behavior in similar ways. In an evolutionary perspective, emotions and feelings are best understood as *modes of functioning* of the brain that coordinate physiological, cognitive, and behavioral responses to fitness-relevant events and situations. In other words, they are superordinate psychological mechanisms designed to coordinate the activity of other mechanisms, including those involved in attention, perception, memory, learning, and physiological regulation (Al-Shawaf et al., 2016; Nesse, 1990, 2004a; Nesse & Ellsworth, 2009; Tooby & Cosmides, 1990, 2008).

To illustrate, fear activates immediate physiological changes in circulation, respiration, and digestion; increases perceptual acuity and improves the quick detection of threats; increases the salience of conceptual categories such as “dangerous” and “safe” (e.g., a closet may suddenly be perceived as a hiding place); directs memory to threat- and escape-relevant information; activates powerful learning systems (fear conditioning); makes some behavioral responses (e.g., running) more likely; and, depending on context, may result in facial and vocal expressions that communicate the presence of a threat to others (Tooby & Cosmides, 2008). Accordingly, affective experiences are associated with the activation of multiple brain regions, with considerable overlap between different emotions (Lindquist et al., 2012; more on this later). Depending on their function, emotions may or may not have specific nonverbal markers (e.g., a characteristic facial expression). While clearly expressing anger may be adaptive, other emotions such as jealousy may be strategically designed to have no obvious external cues (Al-Shawaf et al., 2016; Sznycer et al., 2017).

An important function of some emotions is to recalibrate the value of motivational regulatory variables, both in the affected individual and in others. Intense fear in response to a life-endangering threat may switch the internal representation of the environment from safe to unsafe. Expressing anger in response to a slight may increase the implicit “weight” that others assign to the welfare of the angry individual, making it more likely that the latter will be treated fairly in the future. More generally, emotional expressions elicit emotional states in others, which then feed back on the individual’s own emotions in a loop of reciprocal regulation (Gilbert, 2015; Tooby & Cosmides, 2008, 2015; Tooby et al., 2008). Given their central role in changing and updating internal representations, it is not surprising that emotions can be exploited to manipulate others, either by producing simulated or exaggerated emotional displays (e.g., pretending to be angry or hurt) or by strategically eliciting affective states (e.g., inducing guilt; Al-Shawaf et al., 2016).

EMOTIONS AND MOTIVATIONAL SYSTEMS

Motivational systems control goal-directed behavior by deploying adaptive emotional programs in response to cues and events that activate the system, indicate success or progress toward the system’s proximate goal, or indicate failure or obstacles to achieving the goal. For example, a sudden scream in the dark is likely to activate the fear system, whose goal is to avoid or escape

an immediate threat. The system's activation is marked by the emotion of fear; in turn, fear orchestrates a complex suite of responses, possibly including escape behaviors such as hiding or running away. If one manages to avoid the threat (motivational success), fear will subside and be replaced by relief and a feeling of security; after a while, the fear system will deactivate and give way to other motivational priorities (e.g., hunger). If the threat cannot be avoided (motivational failure), the system may switch from fear to despair, trigger extreme defensive programs such as fainting, or activate the aggression motivational system in an attempt to fight back the threat.

In sum, motivational systems set the goals of behavior and monitor cues of success and failure in the pursuit of those goals. Motivational systems affect behavior indirectly by triggering emotions, which in turn coordinate adaptive suites of physiological, cognitive, and behavioral responses. Different motivational systems may deploy similar or identical emotions in the service of different goals. Anger may be triggered by the aggression system in response to a physical threat, by the attachment system in response to separation from a caregiver, or by the pair bonding system in response to a partner's infidelity (see later discussion for details). Conversely, it is quite possible that emotion terms such as "anger" and "love" actually refer to families of related but motivationally distinct mechanisms. Romantic love has many aspects in common with parental love, has likely evolved out of it, and depends on some of the same psychological and physiological mechanisms (Crespi, 2016a; Esch & Stefano, 2005; Feldman, 2017; Panksepp, 1998). Still, the two differ in important respects and have been shaped by selection to serve different adaptive goals. While aggressive anger overlaps with attachment-related anger well enough that the same word may be used to describe both, the two may differ in subtle ways and be mediated by partially distinct mechanisms.

The intricate relationship between motivation and emotion has two important implications. First, most affective labels have low motivational specificity—knowing that a person feels angry or ashamed says relatively little about his or her goals and motivational state. Second, emotions come in functionally related families that reflect the underlying structure of motivational systems. To the extent that motivational systems are functionally related to one another and/or rely on similar affective mechanisms, the phenotypic map of emotions can be expected to contain

many fuzzy, overlapping clusters rather than a small number of sharply defined “basic” categories (Nesse, 2004a; Nesse & Ellsworth, 2009; Sznycer et al., 2017). In this perspective, there is also no clear-cut distinction between short-lived emotions and long-lasting moods. However, while emotions are tied to the action of specific motivational systems, moods are more likely the product of regulatory processes that affect multiple systems at once.

Human Motivational Systems

Table 2.1 shows a nonexhaustive list of motivational systems in humans (see Aunger & Curtis, 2013; Gilbert, 1995, 2015; Keltner et al., 2006; Kenrick et al., 2010; McGuire & Troisi, 1998; Panksepp, 1998, 2011; Toronchuk & Ellis, 2013). Each system has one or more proximate goals and a characteristic set of emotions that correspond to its activation, success, and failure. For brevity, the list does not include the many feeling-regulated systems that mediate basic physiological needs such as eating, drinking, sleeping, and temperature regulation. Some scholars have argued for a small core of motivational/emotional systems that are phylogenetically ancient and exist in homologous forms across mammalian species (e.g., Panksepp, 1998, 2011). While these criteria apply to some important systems such as fear and attachment, there is no a priori reason to restrict one’s attention to mechanisms that are shared by all mammals. Each species faces its own set of ecological and social challenges, which in turn drive the evolution of specialized motivational systems and goals. Among other things, the human adaptive complex includes extended cooperation with non-kin, extensive reliance on technology and information transmission, pair bonding, and multigeneration transfer of resources. These traits have been critical to human fitness for thousands of generations and have shaped their own sets of motivations and associated emotions such as jealousy, greed, and curiosity (Al-Shawaf et al., 2016; Aunger & Curtis, 2013).

Table 2.1 A Partial List of Human Motivational Systems

Motivational system	Proximate goals	Examples of associated emotions/feelings
Aggression	Counter immediate threats by	<i>Activation:</i> anger, rage,

system	threatening or inflicting harm Harm and/or control other individuals/organisms (e.g., prey)	excitement <i>Success:</i> relief, satisfaction <i>Failure:</i> frustration, fear
Fear system	Avoid or escape immediate threats	<i>Activation:</i> fear, freezing (attentive immobility), panic <i>Success:</i> relief, safety <i>Failure:</i> shock, despair, tonic immobility
Security system	Detect and prevent potential threats	<i>Activation:</i> anxiety, wariness, apprehension <i>Success:</i> relief, safety <i>Failure:</i> fear, guilt
Disgust system	Avoid contact with pathogens and/or toxins (pathogen disgust) Avoid undesirable sexual contacts (sexual disgust) Condemn moral violations (moral disgust)	<i>Activation:</i> disgust, repulsion, nausea <i>Success:</i> relief <i>Failure:</i> discomfort, shame, guilt
Status system	Improve, maintain, and display status (dominance and/or prestige) Defer/submit to higher status individuals	<i>Activation:</i> ambition, anxiety, envy, embarrassment <i>Success:</i> confidence, pride, admiration, awe <i>Failure:</i> frustration, shame, anger, sadness

Mating system	Find, attract, and select sexual partners Engage in sexual intercourse Conceive offspring	<i>Activation:</i> excitement, arousal, desire, embarrassment <i>Success:</i> pleasure, joy, confidence, pride <i>Failure:</i> frustration, sadness, anxiety, shame
Attachment system	Maintain proximity and availability of caregivers	<i>Activation:</i> loneliness, anxiety, fear, anger <i>Success:</i> joy, safety, relief, comfort <i>Failure:</i> sadness, despair, loneliness
Caregiving system	Protect and provision offspring and other dependent individuals	<i>Activation:</i> solicitude, compassion, anxiety <i>Success:</i> joy, satisfaction, pride <i>Failure:</i> sadness, guilt
Pair bonding system	Form and maintain long-term relationships with sexual partners Prevent and/or detect infidelity	<i>Activation:</i> romantic love, jealousy <i>Success:</i> joy, security <i>Failure:</i> sadness, loneliness, anger
Affiliation system	Form and maintain relationships with group members (e.g., friends)	<i>Activation:</i> sympathy, loneliness, anxiety, fear <i>Success:</i> joy, security, belonging

		<i>Failure:</i> sadness, loneliness
Reciprocity system	Find and select cooperation partners Cooperate, trade favors and resources Monitor and enforce fairness	<i>Activation:</i> trust, suspiciousness, envy, indignation <i>Success:</i> joy, satisfaction, pride, gratitude <i>Failure:</i> sadness, anger, guilt
Acquisition system	Acquire, accumulate, and defend resources	<i>Activation:</i> desire, greed, envy, anxiety <i>Success:</i> joy, satisfaction, pride, security <i>Failure:</i> frustration, sadness, anger
Play system	Acquire and improve skills (physical, social, cognitive) Display skills and other attractive traits Promote social bonding	<i>Activation:</i> excitement, engagement <i>Success:</i> joy, confidence <i>Failure:</i> frustration, boredom
Curiosity system	Acquire knowledge Explore the environment	<i>Activation:</i> curiosity, excitement, surprise <i>Success:</i> satisfaction, pride, confidence, insight <i>Failure:</i> frustration, boredom, anxiety, confusion

AGGRESSION SYSTEM

Aggression is a basic motivation to harm or threaten other organisms, including—but not limited to—individuals of the same species. Aggression is often deployed as a defensive strategy by immediate threats to oneself, one's kin, or one's allies (*reactive* aggression). Reactive aggression is marked by intense arousal, anger, or rage, and can be triggered by high levels of fear (Panksepp, 1998, 2011). For this reason, aggression and fear are sometimes discussed together as part of a unitary “fight-or-flight” or “fight-flight-freeze” system (e.g., Corr, 2013). However, aggressive motivations are not necessarily defensive; a prime example of *proactive* aggression is hunting, which involves extreme aggression toward prey but no anger. Proactive aggression is instrumental and cold-blooded, and may even be accompanied by feelings of pleasure and excitement (Elbert et al., 2010; Panksepp, 1998). In humans, proactive aggression is also a key component of group conflicts and wars, in the course of which the enemy is dehumanized and effectively treated like prey (Wrangham, 1999). Proactive aggression can be employed to reinforce dominance hierarchies, take or steal resources, and more generally control the behavior of other individuals. Importantly, humans can use language to make threats, hurt another individual's reputation, exclude him or her from the group, incite others to violence, and so on; thus, *verbal* and *relational* aggression play an important role in our species alongside physical aggression (Archer, 2009).

FEAR SYSTEM

The fear system is an ancient defensive mechanism that motivates organisms to avoid or escape immediate threats. This system can be activated by a multitude of cues and situations, and many specific fears are acquired through learning. However, some types of stimuli elicit fear with no need for learning (e.g., sudden loud noises) or after minimal exposure (e.g., snakes, spiders, angry male faces; LoBue & Rakison, 2013; Mallan et al., 2013; Öhman, 2009). The existence of so-called “innate fears” is highly adaptive; when even a single exposure to a certain recurring threat can lead to death, trial-and-error learning is simply too dangerous to be an option. *Tonic* and *attentive immobility* are important components of the fear system. Attentive immobility or “freezing” occurs in preparation for escape or fighting; tonic immobility is a kind of paralysis or fainting without loss of consciousness, a last resort defense when harm is inevitable. Freezing and tonic immobility

can be associated with analgesia, emotional numbing, and dissociative experiences (Hagenaars et al., 2014; Nijenhuis et al., 1998).

SECURITY SYSTEM

Like the fear system, the security system is an ancient mechanism designed to protect organisms from threats. The crucial difference is that fear is triggered by immediate threats, whereas security motivation is activated by *potential* threats—that is, threats that are relatively rare and difficult to detect but may have catastrophic consequences, such as predators or contaminating pathogens (Boyer & Liénard, 2006; Woody & Szechtman, 2011). Immediate threats evoke fear and escape/fight behaviors; in contrast, activation of the security system is marked by anxiety, wariness, and precautionary repetitive behaviors such as checking and exploration, which are designed to gather further information about the presence of potential risks. Indeed, the security system tends to inhibit fear, preventing flight/panic responses to permit cautious exploration (Graeff, 2004). By its nature, the security system is activated by subtle and indirect cues of danger; once activated, is not easily switched off because the *absence* of a potential threat is hard or even impossible to determine with certainty, and there are no clear signals indicating whether precautionary behaviors have been successful. In humans, security motivation often involves heightened perceptions of responsibility, and failing to avert a preventable threat may lead to painful feelings of guilt.

DISGUST SYSTEM

The disgust system is a defensive mechanism whose original function is preventing contact with pathogens and/or toxins through ingestion of contaminated foods, drinks, or waste products; manipulation of contaminated objects; and contact with infected people (*pathogen* disgust; Curtis, 2011; Toronchuk & Ellis, 2013). Pathogen disgust promotes physical avoidance, expulsion (e.g., vomiting), and cleaning behaviors. Disgust can also trigger activation of the security system; indeed, the two systems often work in synergy. Over evolutionary history, the disgust system has been coopted and differentiated to deal with other kinds of threats (Rozin et al., 2008; Tybur et al., 2013). In particular, *sexual* disgust is designed to prevent sexual contact with partners that would be detrimental to fitness, for example because they are too old, too genetically similar (e.g., siblings and other close kin), or prone to sexually transmitted diseases (e.g., highly promiscuous individuals). Finally, disgust in our species is deeply connected to morality: violations of

moral norms and taboos can elicit disgust and feelings of uncleanness and contamination (e.g., Chapman et al., 2009). A likely function of *moral* disgust is to motivate and coordinate social distancing from individuals who violate moral rules (Tybur et al., 2013). While failure to avoid contact with disgusting objects leads to intense physical discomfort, motivational failures in the sexual and moral domains may evoke self-evaluative emotions such as shame and guilt.

STATUS SYSTEM

Dominance hierarchies are extremely common in animals, and so is the existence of specialized motivational systems designed to deal with the challenges of social rank. Dominance systems have a dual function: that of enhancing or defending one's rank and that of submitting to higher ranking individuals to avoid punishment and retaliation (McGuire & Troisi, 1998; Toronchuk & Ellis, 2013). In our species, social hierarchies reflect both physical dominance and skill-based prestige; the more general concept of a "status system" covers both aspects, emphasizing the complex nature of human competition (Anderson et al., 2015; Aunger & Curtis, 2013; see also Keltner et al., 2006). The status system is activated by challenges to one's dominance rank or prestige (from provocations and disrespectful acts to situations that involve social judgments), but also by opportunities to rise in the social hierarchy. The emotions associated with the system's success include pride and confidence, whereas failure may trigger shame, frustration, anger, and sadness. Importantly, voluntary deference to high-status individuals such as leaders and teachers can evoke a range of positive emotions such as admiration and awe (Keltner et al., 2006). Dominance competition often elicits aggression, and the two systems are deeply connected on a functional level (Anderson et al., 2015; Toronchuk & Ellis, 2013).

MATING SYSTEM

The mating system plays a critical role in reproduction by motivating sexual behavior, from courtship and mate choice to copulation. The system is activated by the presence or prospect of attractive partners and draws on a broad range of feelings and emotions, from excitement to embarrassment to sexual arousal (Al-Shawaf et al., 2016; McGuire & Troisi, 1998; Panksepp, 1998). Ultimately, the function of the mating system is to conceive offspring; however, its core proximate goals are sexual pleasure and intercourse rather

than conception per se. Sexual desire can thus become largely decoupled from reproduction, for example as a result of effective birth control. At the same time, there are some indications that a conscious desire to have children may be an evolved feature of mating motivation in humans, particularly in women (Brase & Brase, 2012; Rotkirch, 2007; Rotkirch et al., 2011).

ATTACHMENT SYSTEM

Like most young mammals, infants and children are vulnerable and depend on adults for feeding and protection. The attachment system is designed to monitor and maintain the proximity and availability of caregivers; physical separations quickly activate the system, as do perceptions of danger, pain, and discomfort. Emotional responses to separation include anxiety and angry protests, whereas a sensitive response from the caregiver reliably deactivates the system and evokes feelings of safety and relief. In infancy and childhood, attachment has high motivational priority, consistent with its critical role in ensuring survival. When activated, the attachment system inhibits play and curiosity; conversely, the presence of an available attachment figure works as a “secure base” for exploration (Kobak & Madsen, 2008; Marvin & Britner, 2008; Posada et al., 2013). In all likelihood, the mammalian attachment system originally evolved from mating-related mechanisms promoting physical proximity and contact. In our species, attachment has been coopted as a building block of close relationships in adulthood, including those with romantic partners and friends (Crespi, 2016a; Del Giudice, 2009a; Fletcher et al., 2015; Panksepp, 1998). Specifically, most intimate relationships involve an attachment component as they provide comfort, reassurance, and safety in times of distress.

CAREGIVING SYSTEM

Mirroring the biological function of the attachment system, the caregiving system motivates parents and other caregivers to protect and nurture their dependent young (Brown et al., 2012; Cassidy, 2008; Panksepp, 1998, 2011). As a species, humans show many features of *cooperative breeding*: across societies, care and protection are provided not just by parents but by multiple individuals including older siblings, grandparents, and friends (Hrdy, 2005; Kramer, 2010). Thus, caregiving motivations are not restricted to one’s biological offspring; conversely, children are predisposed to form multiple attachment relationships in addition to the one with their primary caregiver (usually the mother; Howes & Spieker, 2008). Caregiving is primarily

activated by displays of vulnerability and/or distress (e.g., crying, cute baby-like features) and involves a range of emotions from anxious solicitude to parental love. Failures of the caregiving motivation can trigger powerful negative emotions, including sadness and guilt. Like attachment, caregiving is a vital component of close relationships between adults, particularly intimate friends and romantic partners.

PAIR BONDING SYSTEM

Pair bonding is a central feature of the human adaptive complex. It has likely evolved from the integration of sexual attraction with attachment and caregiving—two motivations rooted in parent–child relations—and reused to enable long-term bonding between sexual partners. In part, romantic love can be seen as a blend of emotions associated with these three systems; at the same time, the psychology of love also shows unique features and evidence of functional specialization. For example, being in love temporarily inhibits the desire for alternative sexual partners, thus working as a “commitment device” in view of shared parental investment (Eastwick, 2009; Fletcher et al., 2015; Quinlan, 2008). Also, romantic love is powerfully associated with jealousy, an emotional mechanism designed to prevent infidelity by partners (Buss, 2000, 2013). For these reasons, it makes sense to postulate a specialized pair bonding system with the specific goal of forming and maintaining couple relationships (Aunger & Curtis, 2013), partly by coordinating other motivational systems such as mating, attachment, caregiving, and reciprocity.

AFFILIATION SYSTEM

Affiliation is a key motivational substrate of group living; its function is to enable and sustain long-term relationships with extended kin and other group members, including friends and allies. Long-term friendships have been documented in many animal species, including a number of primates (Seyfarth & Cheney, 2012). As with pair bonding, the psychology of affiliation significantly overlaps with that of attachment; at the same time, friendship and group membership appear sufficiently distinct from parent–child relations to warrant the hypothesis of a specialized motivational system (Aunger & Curtis, 2013; Bugental, 2000; Kenrick et al., 2010). The affiliation system can be activated by the perception of shared interests and goals, but also by threatening situations or a lack of social resources (isolation, rejection). Successful affiliation evokes feelings of security and

belonging, promotes the formation of a shared group identity, and favors cooperation and reciprocity.

RECIPROCITY SYSTEM

While the affiliation system promotes affective bonding with other group members, the reciprocity system deals with cooperation and with the exchange of favors and resources. Cooperation and reciprocal exchange are greatly facilitated when they occur within families and groups, but they also take place between strangers, or—under certain conditions—even between enemies (Bugental, 2000; Keltner et al., 2006; McGuire & Troisi, 1998). While extensive cooperation networks of non-kin are unique to humans, other primates do engage in more limited forms of reciprocity, for example in the context of grooming and food sharing (Jaeggi et al., 2013; Jaeggi & Gurven, 2013; Engelmann et al., 2015). The reciprocity system can be activated by opportunities such as the presence of a capable and trustworthy partner or by obstacles such as cheating and unfairness. In humans, reciprocity is supported by specialized cognitive mechanisms that monitor violations of rules and keep track of partners' contributions and reputations over time (Cosmides & Tooby, 2013, 2015). Failures of reciprocity may arouse intense anger or guilt, depending on whether one is the victim or the perpetrator. Both friendships and long-term couple relationships involve elements of reciprocity, although friends and partners are considerably more tolerant than strangers in tracking each other's contributions over time (e.g., Xue & Silk, 2012).

ACQUISITION SYSTEM

An obvious but sometimes overlooked characteristic of our species is the extent to which we store and accumulate resources for future use. Material wealth—whether in the form of land, cattle, houses, or money—provides immediate adaptive benefits as it improves both mating success (especially in men) and the survival of children (Borgerhoff Mulder & Beheim, 2011; Cronk, 1991; Koster, 2011; Nettle & Pollet, 2008). Moreover, stored resources reduce risk by working as a buffer against periods of scarcity and can be passed down from one generation to the next, with cumulative effects on long-term fitness (Borgerhoff Mulder et al., 2009; Winterhalder et al., 1999). Unsurprisingly, humans have strong motivations to acquire resources, accumulate them, and defend them against theft, as well as a distinct psychology of ownership based on emotions such as desire, envy, and greed. The acquisition system likely has its evolutionary roots in the mechanisms

that mediate foraging and food hoarding (Aunger & Curtis, 2013; Preston, 2014; Preston & Vickers, 2014). The specific goals of the acquisition system depend on its interaction with other motivations such as mating and pair bonding. For example, saving resources for future family needs in the context of a long-term relationship is not the same as acquiring costly luxury goods to boost success in courtship and short-term mating (Griskevicius et al., 2007; Sundie et al., 2011).

PLAY SYSTEM

Play is virtually universal in mammals and absorbs a large fraction of juveniles' time and energy. The overarching function of play is to enable self-training in a range of adaptive skills; fighting, parenting, and foraging are prominent recurring themes across species. More specific functions are regulating neuromuscular development and learning how to cope with unexpected events (Burghardt, 2005; Byers & Walker, 1995; Spinka et al., 2001). While play motivation clearly constitutes an independent psychological mechanism (Panksepp, 1998), it always works in synergy with other motivational systems that provide the momentary goals of play and the relevant behavioral/emotional repertoires. For example, rough-and tumble play involves the playful coordination of the aggression, fear, and status systems. Cognitive skills are also exercised through play, as when children play games of memory, numbers, and language (e.g., Locke & Bogin, 2006). Finally, play promotes social bonding in synergy with the affiliation system (Toronchuk & Ellis, 2013) and can be an effective way to *display* skills and other attractive traits (e.g., physical strength, intelligence) to potential allies and future partners. A notable aspect of play in humans is that adults also play, not only with children but also with other adults. Bonding and displaying are likely to be the main functions of adult play, compared with the primary training role of play in childhood.

CURIOSITY SYSTEM

Curiosity is the motivation to acquire knowledge and has been long recognized as a fundamental motive in animals (Aunger & Curtis, 2013; Loewenstein, 1994). Information-seeking is essential to build models of the world and improve one's ability to make inferences and predictions (Gottlieb et al., 2013). In humans, knowledge can be used to build social prestige or increase one's value as a social partner (Dessalles, 2007). Far from being a "cold" cognitive task, the acquisition of information is regulated by a wide

range of emotions and feelings, from excitement and surprise to boredom, frustration, and anxiety. Curiosity and exploration are often discussed in association with play, and play is certainly a powerful way to gather information about oneself, other people, and the environment. Even pretend play based on unrealistic scenarios can play a critical role in building sophisticated causal models of the world (Weisberg & Gopnik, 2013). However, there are many ways to acquire knowledge that do not rely on play; moreover, language permits massive transfer of information without the need for first-hand experience. In humans, adaptations for learning seem to be matched by adaptations for *teaching*, the deliberate transmission of knowledge and skills (Csibra & Gergely, 2006).

Defensive Mechanisms and the Smoke Detector Principle

Several of the mechanisms described in the previous section—most notably the fear, security, and disgust systems—can be understood as defenses whose function is to protect an individual from physical or social damage. The compromises involved in the design of defensive mechanisms have important implications for behavior, which are encapsulated by the *smoke detector principle* (Nesse, 2001a, 2005a). By necessity, the calibration of defenses involves a tradeoff between the rate of false negatives (failing to activate a defense mechanism when a threat is present) and that of false positives or “false alarms” (mistakenly activating the mechanism when no threat is present). A more sensitive system—for example one with a lower threshold for experiencing disgust—offers enhanced protection from false negatives (e.g., accidentally eating contaminated food) but does so at the cost of an increased risk of false positives (e.g., rejecting safe food items). It is often the case that false negatives have potentially catastrophic consequences, while false positives have only minor costs. Under such conditions, natural selection favors a high rate of false positives, leading to the evolution of hair-trigger defenses that often activate when no actual threat is present (Figure 2.1).

The smoke detector principle applies to all sorts of defensive mechanisms, including physiological defenses such as fever and vomiting (Nesse, 2001a). The principle has two key implications. First, high rates of errors and inappropriate responding may be adaptive features of a system rather than

signs of dysfunction. Second, blocking or attenuating an adaptive defense (e.g., taking fever-reducing or anxiety-reducing drugs) may have little or no adverse effects most of the times, precisely because many instances of activation are likely to be false positives rather than necessary responses (Nesse & Jackson, 2006).

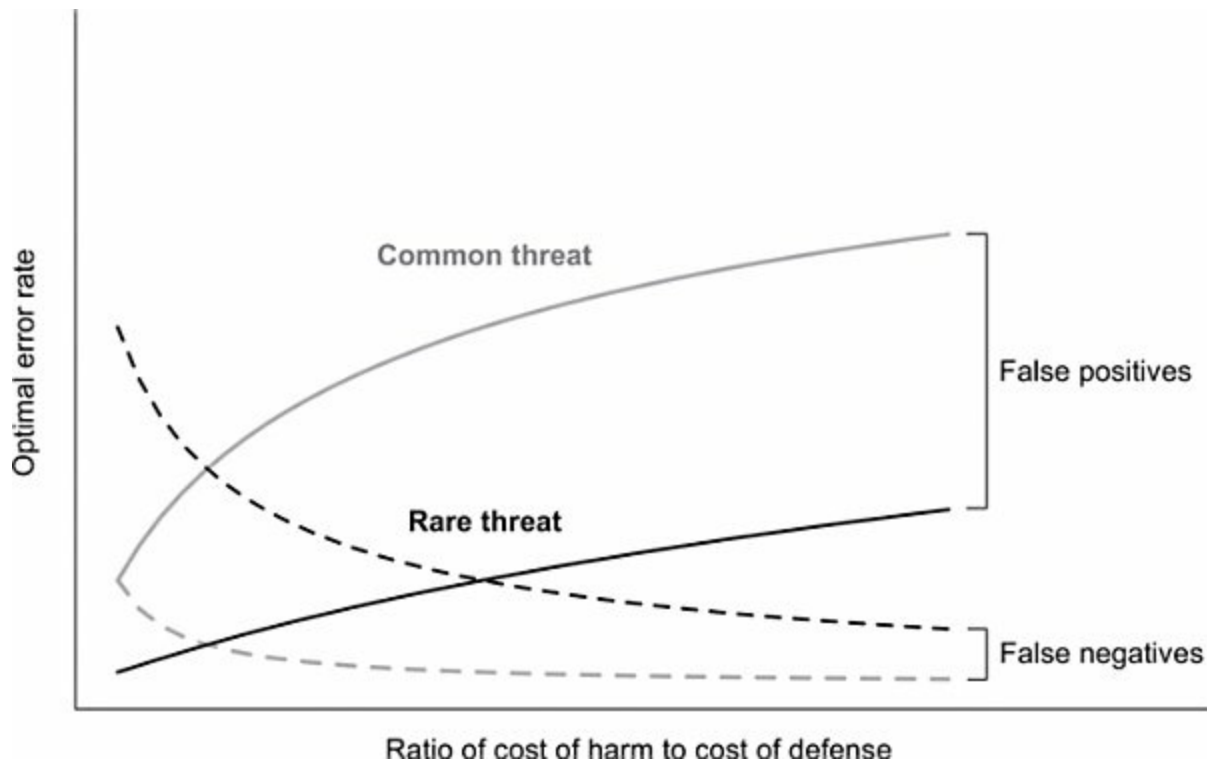


Figure 2.1. The smoke detector principle in the calibration of defensive mechanisms. When threats are common (gray line) and/or the fitness cost of potential harm is large relative to the cost of the defense, it is adaptive to accept a high rate of false positives (activating the defense even if no threat is present) in order to reduce the rate of false negatives (failing to activate the defense when a threat is present).

Motivation and Emotion in the Brain

Motivation is the engine of the mind, and motivational systems coordinate brain activity at all levels, from the brainstem to the cortex. Studies of motivation and emotion show an important role for the thalamus, hypothalamus, and other subcortical structures such as the amygdala, the bed nucleus of the stria terminalis (BNST), the periaqueductal gray (PAG), and the ventral striatum (including the nucleus accumbens, NAcc). The amygdala is a central node in this network (often referred to as the *limbic system*) and

plays a key role in encoding salience and orienting attention to motivationally relevant stimuli. The main cortical areas involved in affective processes are the medial temporal lobe (including the hippocampus and parahippocampal gyrus), the insula, the anterior and posterior cingulate cortex (ACC and PCC), the dorsomedial prefrontal cortex (DMPFC), and the ventromedial prefrontal cortex (VMPFC, including the orbitofrontal cortex, OFC; see Lindquist et al., 2012; Panksepp, 2011; Toronchuk & Ellis, 2013). A map of these brain regions is shown in [Figure 2.2](#). Acronyms for the brain regions discussed in the book are listed in [Box 2.1](#).

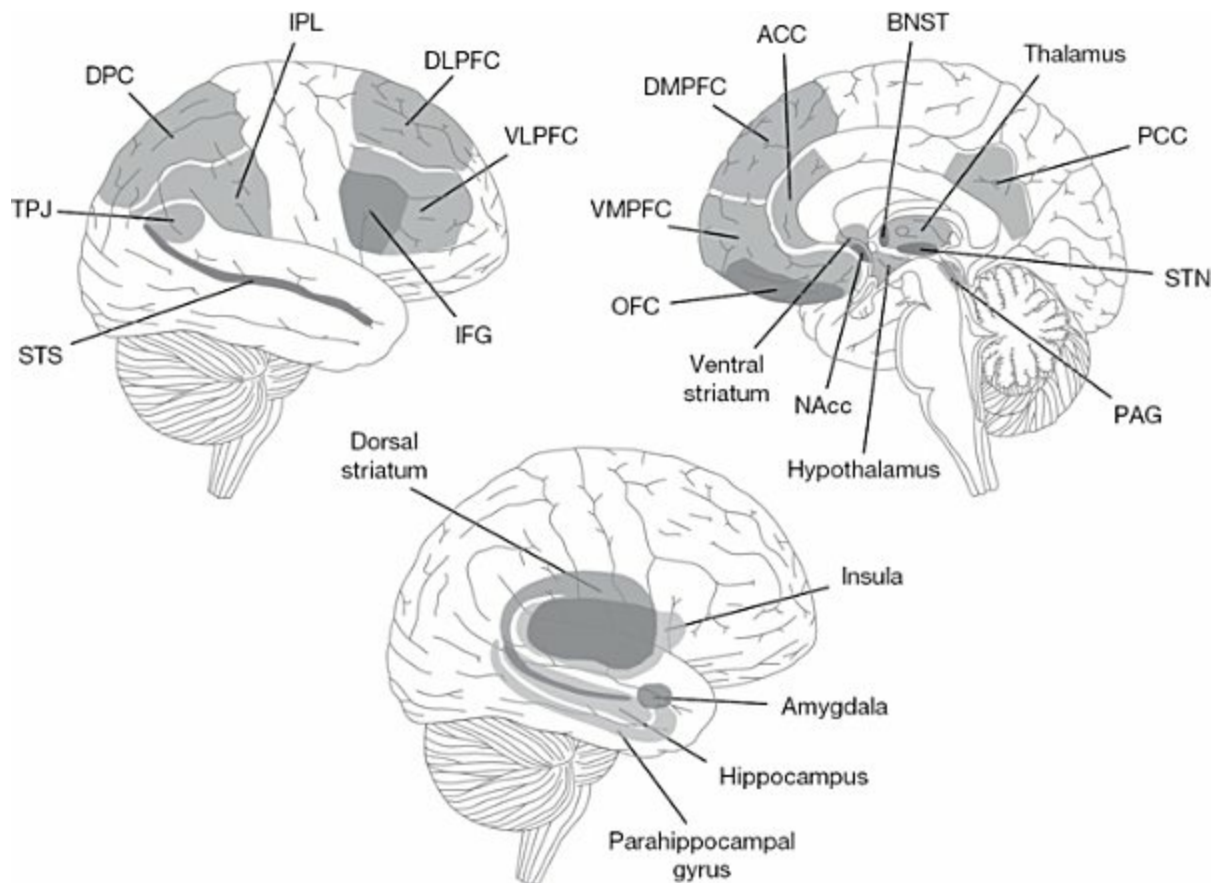


Figure 2.2. Map of the brain structures discussed in this chapter. ACC = anterior cingulate cortex. BNST = bed nucleus of the stria terminalis. DLPFC = dorsolateral prefrontal cortex. DMPFC = dorsomedial prefrontal cortex. DPC = dorsal parietal cortex. IFG = inferior frontal gyrus. IPL = inferior parietal lobule. NAcc = nucleus accumbens. OFC = orbitofrontal cortex. PCC = posterior cingulate cortex. STN = subthalamic nucleus. STS = superior temporal sulcus. TPJ = temporoparietal junction. VLPFC = ventrolateral prefrontal cortex. VMPFC = ventromedial prefrontal cortex.

Box 2.1

Acronyms for Brain Regions and Structures

ACC: anterior cingulate cortex
BNST: bed nucleus of the stria terminalis
DLPFC: dorsolateral prefrontal cortex
DMPFC: dorsomedial prefrontal cortex
DPFC: dorsal prefrontal cortex
DPC: dorsal parietal cortex
IFG: inferior frontal gyrus
IPL: inferior parietal lobule
LC: locus coeruleus
MPFC: medial prefrontal cortex
NAcc: nucleus accumbens
OFC: orbitofrontal cortex
PAG: periaqueductal gray
PCC: posterior cingulate cortex
PVN: paraventricular nucleus
SN: substantia nigra
SON: supraoptic nucleus
STN: subthalamic nucleus
STS: superior temporal sulcus
TPJ: temporoparietal junction
VLPFC: ventrolateral prefrontal cortex
VMPFC: ventromedial prefrontal cortex
VTA: ventral tegmental area

DECISION-MAKING AND SELF-REGULATION

Decision-making and self-regulation are intrinsic features of goal-directed behavior. Virtually all of an organism's actions can be described as choices between potential alternatives; the organism's biological goals determine which actions should be favored or avoided and how their outcomes should be evaluated. In other words, the distinction between motivation, decision-making, and self-regulation is largely artificial—they are all facets of the general problem of adaptive behavior control. Still, the distinction makes it

easier to look at the problem from complementary perspectives and facilitates discussion of some important concepts.

Decision-Making

BAYESIAN OPTIMALITY AND THE ROLE OF HEURISTICS

Adaptive behavior requires organisms to make approximately optimal (or at least “good enough”) decisions given their own state and that of the environment. However, organisms cannot know the environment directly but only imperfectly through their senses; moreover, conditions are always changing in more or less predictable ways, and the outcomes of actions cannot be anticipated with certainty. In a stochastic world filled with uncertainty and noise, decision-making crucially depends on the ability to estimate and combine probabilities. When making inferences and predictions, current sources of information must be integrated with previous expectations, which in turn may result from learning, evolved predispositions, or both. The optimal way to combine probabilistic information is described by the equation known as *Bayes’ rule*:

$$P(A|B) = \frac{P(A)P(B|A)}{P(B)}$$

In a nutshell, the rule describes how the probability of event A (e.g., the probability that it will rain today) changes after observing event B (e.g., clouds in the sky). $P(A)$ is the previous expectation or *prior* (in the example, the probability of rain on any given day), while $P(A|B)$ is the conditional probability of A given that B has occurred—that is, an updated or *posterior* probability based on all the available evidence. To obtain the posterior probability, the prior is multiplied by the conditional probability of B given A (the probability of observing clouds on a rainy day) and divided by the overall probability of B (the probability of observing clouds averaged across rainy and nonrainy days). From another perspective, the logic of Bayes’ rule consists of using prior expectations to make a prediction about the state of the world, then updating one’s expectations based on the discrepancy between the predicted and observed state (*prediction error*). Bayes’ rule is the

foundation of an expanding family of models that cover all kinds of psychological processes, from perception and memory to motor and cognitive control (Clark, 2013; Dayan & Daw, 2008; Jacobs & Kruschke, 2011; Jiang et al., 2014; Mathys et al., 2011; Vilares & Körding, 2011).

Given that optimal decision-making must conform to Bayes' rule, the IMA principle implies that organisms will generally tend to behave *as if* they were performing Bayesian computations to maximize their expected fitness. In this sense, adaptation and Bayesian decision-making are two sides of the same coin (Geisler & Diehl, 2003; Trimmer & Houston, 2014). Notably, the smoke detector principle described in the previous section is a straightforward application of Bayes' rule to the optimal calibration of defenses. However, actual psychological mechanisms do not work by literally applying Bayes' rule to probabilities, not least because full-fledged Bayesian computations for realistic problems are too complex and time-consuming to be solved in real time. Instead, selection favors the evolution of mechanisms that employ Bayesian approximations or even “rules of thumb” (*heuristics*) that mimic the results of Bayesian logic by performing much simpler operations, often relying on limited amounts of information (Castellano, 2015; Lange & Dukas, 2009; Mathys et al., 2011; Sanborn & Chater, 2016; Trimmer & Houston, 2014). On the positive side, heuristics can be fast and cognitively efficient; precisely because of their simplicity, they also tend to be very robust to perturbations in the environment, sensory and processing noise, and unreliable information (Gigerenzer, 2008; Gigerenzer & Brighton, 2009). At the same time, heuristics usually work by making implicit assumptions about the environment; a heuristic that produces near-optimal behavior in a certain context may fail spectacularly if conditions change. From this vantage point, emotions and motivational systems can be framed as sophisticated heuristic devices. They embody assumptions (for example about pathogens, children, or status hierarchies) that have been shaped and refined in the EEA and coordinate fast behavioral responses that are likely—but not guaranteed—to be adaptive in most situations (Johnson-Laird & Oatley, 1992).

TRADEOFFS IN DECISION-MAKING

Even well-adapted decision-makers face a number of unavoidable tradeoffs. A ubiquitous tradeoff is that between *speed* and *accuracy* (Chittka et al., 2009; Keramati et al., 2011; Trimmer et al., 2008). Waiting to collect more information from the environment reduces the margin of error for a decision; at the same time it delays action, which can be costly for the organism.

Speed–accuracy tradeoffs also occur when engaging in more detailed computations (e.g., a systematic exploration of all the options instead of a quick decisional heuristic) would permit better decisions but also take longer to complete. Similar tradeoffs involve the flexibility of a system, or its ability to effectively handle a variety of inputs and situations. While flexibility can be useful, it usually comes at a cost in terms of other desirable qualities such as speed and robustness. For example, learning mechanisms that build a detailed internal model of the environment (instead of simply storing stimulus–response associations or other summary representations) provide enhanced flexibility and accuracy; however, they also require complex and lengthy computations to make decisions, and are more sensitive to the effects of noise in the processing machinery (Daw & Dayan, 2014; Daw et al., 2005; Dayan & Daw, 2008).

In many cases, future decision-making can be improved by spending time and energy to explore the environment and try out new behaviors. Of course, the time spent exploring and learning cannot be used to pursue rewards and thus comes at an immediate cost for the organism. This is the essence of the *exploration–exploitation* tradeoff (Bland & Schaefer, 2012; Dayan & Daw, 2008). Exploration–exploitation tradeoffs may occur at different time scales. An animal searching for food in a new environment faces a short-term tradeoff between eating the first items it finds and exploring in the hope of finding better or more abundant food. On a developmental scale, the time that young animals spend playing and learning represents a long-term investment in exploration at the expense of immediate exploitation (Berger-Tal et al., 2014; Eliassen et al., 2007).

RISK SENSITIVITY AND TIME DISCOUNTING

Decision-makers deal with risk whenever the result of an action cannot be predicted with certainty. In biology and economics, risk is defined as *unpredictable variation in outcomes* or “expected uncertainty” (Bland & Schaefer, 2012; Kacelnik & Bateson, 1997; Smallwood, 1996). Whereas some actions offer a narrow range of possible outcomes (low risk), others entail widely variable outcomes (high risk), with the potential for large gains as well as large losses. For example, hunting large but hard-to-catch prey is riskier than feeding on smaller but easily caught animals. Risk-sensitive organisms do not only weigh expected outcomes, but also choose among options based on outcome variability.

Whether natural selection favors positive or negative attitudes toward risk

crucially depends on the shape of the relationship between outcomes and fitness (Figure 2.3). When better outcomes yield diminishing fitness returns (e.g., a satiated animal benefits less from eating larger amounts of food), a smaller variance in outcomes is associated with higher expected fitness, and *risk aversion* becomes optimal. Since riskier actions often yield larger expected rewards, risk-averse organisms routinely face *risk–reward* tradeoffs (Sih & Del Giudice, 2012). When better outcomes are associated with increasing returns (e.g., a starving animal would benefit disproportionately more by eating a larger amount of food), the expected fitness is maximized by choosing options with higher variance, and *risk proneness* is favored by selection. Risk proneness puts a premium on variance, so that risky options can be preferred even if they have a worse expected outcome. In many real-world scenarios, the fitness consequences of an action are more complex than the simple functions shown in Figure 2.3 and involve multiple thresholds where fitness returns switch from increasing to diminishing and vice versa (Hurly, 2003; Kuznar, 2002; Mishra, 2014). To illustrate, a foraging animal may be risk prone when near starvation (increasing returns), become risk averse once satiated (diminishing returns), but switch to risk proneness again if gaining more energy allows it to reproduce or compete more effectively for status (increasing returns).

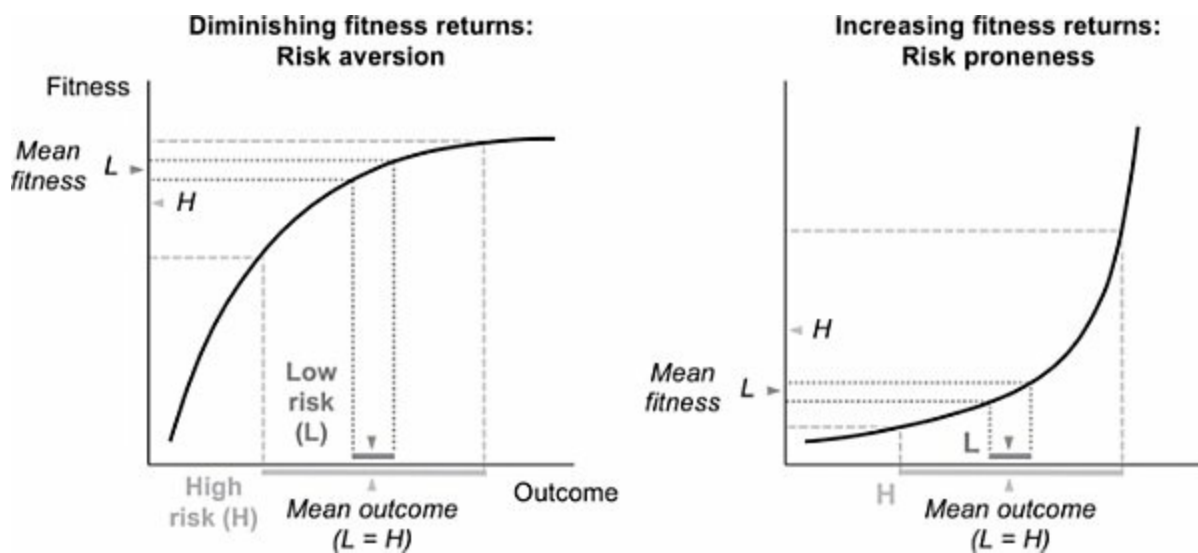


Figure 2.3. The logic of risk sensitivity. When better outcomes yield diminishing fitness returns (concave function on the left), low-risk options with smaller variance have higher mean fitness and risk aversion is favored. With increasing fitness returns (convex function on the right), high-risk options with larger variance have higher mean fitness and risk proneness is favored. Adapted with permission from Frankenhuis & Del Giudice (2012).

In addition to variance, the timing of outcomes plays a critical role in decision-making. It is often the case that larger rewards can be gained by choosing to wait for a longer time. However, waiting for larger rewards also has two downsides: first, conditions may change so that the reward may no longer be available (in the extreme, the organism might die before getting the reward); second, waiting may have opportunity costs as it prevents the organism from engaging in other rewarding activities (Fawcett et al., 2012; Stephens et al., 2004). For these and other reasons, organisms tend to *discount* the future; that is, they subjectively regard delayed gains as less valuable (and delayed losses as less aversive). Moreover, selection often favors a steeper discounting rate for short delays, so that rewards lose most of their subjective value in the initial phase of the waiting period (Fawcett et al., 2012; Green & Myerson, 1996). The key implication of the evolutionary literature on time discounting is that impulsive decision-making (choosing smaller immediate rewards instead of waiting for larger ones) is not necessarily irrational, and may actually be adaptive under a broad range of conditions.

The concept of risk as unpredictable variation does not fully correspond to how risk is usually conceptualized in psychology and the health sciences. “Risky behaviors” such as fighting, stealing, taking drugs, and unprotected sex are primarily defined by a high likelihood of negative outcomes rather than outcome variance per se. However, the same behaviors can often be reframed as involving increased variance—for example, winning a fight may lead to a gain in dominance, whereas losing may bring social defeat in addition to physical harm (Mishra, 2014). More generally, the psychological concept of risk-taking involves a mixture of increased outcome variance and steep discounting of future rewards and/or future harm (e.g., drinking may bring about social advantages in the short term despite its long-term health consequences; see Fox & Tannenbaum, 2011).

Self-Regulation

There are three main ways to think about self-regulation. In the broadest sense, the term describes the control strategies that organisms use to enact their decisions and reach their goals in spite of disturbances such as noise, errors, and unpredictable changes in the environment. More narrowly, self-

regulation refers to a set of processes that do not have well-defined goals such as avoiding threats or finding sexual partners, but rather modulate the activity of multiple motivational systems at once. These mechanisms are useful to coordinate an individual's motivational state across domains or resolve conflicts between incompatible motivations. Finally, some psychological mechanisms support conscious, deliberate forms of self-regulation or self-control. The mechanisms involved in deliberate self-control are usually labeled *executive functions* to distinguish them from more automatic and/or less accessible regulatory processes (Barkley, 2001, 2012; Carver & Scheier, 2011; Del Giudice, 2015a; Goldstein & Naglieri, 2014; Hofmann et al., 2012).

FEEDBACK AND FEEDFORWARD CONTROL

Feedback and *feedforward control* are the basic building blocks of self-regulation, from simple motor coordination all the way up to the complexities of social interactions (Albertos & Mareels, 2010; Bechhoefer, 2005; Del Giudice, 2015a; Wolpert et al., 2003). In feedback or *closed-loop* control, the outcomes of behavior are constantly monitored and used to generate corrective actions, often by reducing the discrepancy between the current state of the system and the desired one (i.e., the goal). Feedback control is a powerful strategy for reaching a desired state even when the outcomes of one's actions cannot be accurately predicted and/or when one lacks a good model of the environment. Feedback loops can be easily nested within one another, creating a hierarchy of goals and subgoals. Nested feedback loops allow a complex control task to be split into simpler, smaller tasks, and permit high levels of flexibility and accuracy (see Carver & Scheier, 2011).

In feedforward or *open-loop* control, an action or sequence of actions is initiated and carried out with no further feedback about the state of the system. Reflexes represent the simplest kind of open-loop control, as in the automatic withdrawal response that occurs on touching a burning object. Indeed, the optimal regulation of defensive mechanisms often involves a rapid initial open-loop response that is then adjusted by slower feedback processes (Shudo et al., 2003). More complex versions of feedforward regulation are based on predictive models and may involve strategic planning, simulation of future events, and integration of existing knowledge about the rules that govern the system and the influence of contextual variables. While feedforward control can be faster and more robust than its feedback counterpart, it requires a considerable amount of contextual

information and a detailed model of the environment and the organism's actions.

Feedback and feedforward strategies have complementary strengths and weaknesses, and selection promotes the integration of both kinds of processes in the mind/brain (e.g., Braver et al., 2009; Wolpert et al., 2003). In the self-regulation literature, this duality is reflected in the distinction between *reactive* and *proactive* control processes. Reactive processes are “late correction” feedback mechanisms that detect and respond to interferences and unexpected events after they occur. Proactive processes anticipate future events and prepare the organism for action based on goal-relevant information; they represent a prime example of model-based feedforward regulation (Braver, 2012; Tops et al., 2010).

APPROACH, AVOIDANCE, AND INHIBITION

Classic theories of motivation posit a fundamental distinction between approach motivations elicited by rewarding stimuli and avoidance motivations elicited by threats and punishments (see Carver & Scheier, 2011; Corr, 2013). In a broader perspective, approach and avoidance can be seen as general regulatory processes that are involved in the operation of most motivational systems. The fear system may trigger either avoidance (flight) or approach behaviors (fight) in different situations; the mating system promotes approach to desirable sexual partner but also avoidance of unattractive ones; and the same food may be rewarding when one is hungry but aversive when one has had too much to eat. There is evidence of specialized psychological mechanisms that regulate the overall balance between approach and avoidance across motivational domains, based on expectations about the future likelihood of rewarding and aversive events; these mechanisms are likely responsible for long-lasting mood states including anxious, euphoric, and depressed moods (Harmon-Jones & Harmon-Jones, 2015; Nettle & Bateson, 2012; Tops et al., 2010; Trimmer et al., 2013).

Reinforcement sensitivity theory (RST; Gray & McNaughton, 2000; Corr, 2013) is a popular approach–avoidance model that postulates the existence of three major systems: a *behavioral approach system* (BAS) that mediates approach to rewards, a *fight-flight-freeze system* (FFFS) that mediates active avoidance of threats and largely overlaps with the fear system described earlier, and a *behavioral inhibition system* (BIS) that suppresses immediate behavioral responses when there is a motivational conflict between the BAS

and the FFFS. In reinforcement sensitivity theory, the BIS is associated with anxiety and promotes exploration or cautious approach to ambiguous stimuli. In all likelihood, the security system described in the previous section derives from basic conflict resolution mechanisms such as the BIS.

The inhibition of behavioral responses is an essential psychological function and is not limited to the specific case of the BIS. Indeed, the brain contains multiple inhibitory pathways that mediate both reactive forms of inhibition (stopping in response to an unexpected event) and proactive ones (preparing to stop or suppress a certain behavior in view of a superordinate goal; Aron, 2011; Diamond, 2013). Inhibition can occur automatically or as the result of a deliberate decision; in fact, the ability to consciously override cognitive and behavioral responses is a central component of self-control.

EXECUTIVE FUNCTIONS

Executive functions are deliberate, self-directed processes that regulate the operation of other psychological mechanisms. The core executive functions are usually described as *inhibition* (the ability to override dominant or prepotent responses), *updating* (the ability to monitor and modify the contents of working memory), and *flexibility* or *shifting* (the ability to switch fluidly and rapidly among tasks, strategies, and contexts). In general, different executive processes tend to reinforce one another—for example, maintaining a goal in working memory makes it easier to inhibit incompatible goals and responses. At the same time, inhibition and flexibility often have opposite effects on task performance, creating the opportunity for executive tradeoffs (Blackwell et al., 2014; Miyake & Friedman, 2012). Planning and problem-solving can be regarded as higher order abilities that depend critically on core executive functions (Diamond, 2013). In a broader perspective, the family of executive functions can be expanded to include self-directed actions such as emotional regulation, internal speech, and mental “play” used to simulate and generate novel action sequences and problem-solving strategies (Barkley, 2001, 2012).

The common theme of executive functions is their orientation toward the future. In different ways, executive processes contribute to detach behavior from immediate, short-term goals in order to maximize the long-term outcomes for the individual. In humans, they facilitate social interactions and provide critical support to social behaviors that extend in time, such as delayed reciprocity, long-term cooperation, and romantic pair bonding. Executive functions are also involved in social competition, because they

enable effective deception as well as self-defense against emotional and cognitive manipulation by others (Barkley, 2001, 2012; Denckla, 1996).

Decision-Making and Self-Regulation in the Brain

Given the functional overlap among motivation, decision-making, and self-regulation, it is not surprising that several brain regions have been found to participate in all these aspects of mental functioning (Figure 2.2). Medial prefrontal areas (VMPFC/OFC), the PCC, and the striatum work together to process rewards and punishments and integrate the motivational value of stimuli with contextual information. These regions are crucially involved in the evaluation of risk and delayed rewards; the hippocampus and amygdala contribute as well, probably because they integrate current stimuli with representations of past and future events and their emotional significance (Aston-Jones & Cohen, 2005; Harmon-Jones & Harmon-Jones, 2015; Moore & Depue, 2016; Pessiglione & Lebreton, 2015; Peters & Büchel, 2011; Rilling, 2011).

The cingulate cortex—particularly the ACC—responds to prediction errors, novelty, risk, and uncertainty, and is one of the main neural correlates of executive functions (Cavanagh & Frank, 2014; Chung et al., 2014). Executive functions also depend on the lateral and medial prefrontal cortex, in combination with the insula and other parietal regions. Specifically, DLPFC activation (particularly in the right hemisphere) is a robust correlate of self-control, risk aversion, and future-oriented decision-making and is implicated in both proactive and reactive control (Braver, 2012; Chung et al., 2014; Knoch & Nash, 2015). The brain contains multiple inhibitory pathways involving lateral and medial prefrontal regions as well as the dorsal striatum. Many of these pathways have a common final node in the subthalamic nucleus (STN). STN activity controls motor inhibition and contributes to regulate the tradeoff between speed and accuracy in decision-making (Aron, 2011; Cavanagh et al., 2011).

MENTALISTIC AND MECHANISTIC COGNITION

From the perspective of agents with minds, the world contains two radically

different kinds of things: mindless physical objects and other agents with minds of their own. The distinction between *mentalist* and *mechanistic* cognition (Badcock, 2004; Crespi & Badcock, 2008) crisply differentiates between two types of psychological mechanisms: those designed to interact with physical objects and those specialized to interact with other minds. The separation between mentalistic and mechanistic cognition is far from complete, since many cognitive processes (e.g., attentional orienting, reinforcement learning, cause–effect reasoning) can be applied to both social and nonsocial information (Heyes, 2014; Parkinson & Wheatley, 2015; Schaafsma et al., 2015). At the same time, there is considerable evidence of functional specialization in the mechanisms that mediate our interactions with the “world of things” on the one hand and the “world of minds” on the other.

Mentalistic Cognition

Mentalistic cognition—also known as *mentalizing*, *theory of mind*, and *folk psychology*—is an umbrella term for the psychological mechanisms that allow us to infer, represent, and respond to the mental states of other individuals. The attribution of mental states to another agent has been called the *intentional stance* (Dennett, 1971). In essence, the intentional stance is an evolved heuristic for predicting behavior using a simplified model of the agent’s mind based on intuitive concepts like beliefs, desires, and emotions; this is incomparably more powerful and efficient than attempting to use a detailed physical model of the agent to predict its behavior. Mentalistic processes do not necessarily depend on explicit or conscious representations of others’ mental states. For example, while the ability to follow gaze does not require a representation of the gazing individual’s perceptual and attentional states, it is designed to operate on the assumption that there are certain systematic links between gaze direction, attention, and perception.

Mentalizing is best described as a complex mosaic of mechanisms rather than a single ability (Barrett, 2015; Frith & Frith, 2010; Schaafsma et al., 2015). Some of those mechanisms are automatic, others deliberate; some are strongly connected to emotional processes, while others depend on cause–effect reasoning; some are evolutionarily ancient, while others have evolved only recently. Whereas some components of mentalistic cognition are highly specialized for mentalistic inference (e.g., recognition of emotional

expressions), others have a broader range of functions (e.g., face recognition); finally, some mechanisms may support or replace mentalizing in daily life without being specifically designed to process social information (e.g., attentional orienting; Heyes, 2014). A strong case can be made that mentalizing mechanisms originally evolved to predict the behavior of other people and were later adapted to represent and interpret one's own mental states; self-directed mind-reading or *metacognition* dramatically increases the ability to regulate one's cognitive processes—thus overlapping with executive functions—and communicate them to other people (Carruthers, 2009).

MENTALIZING AND MOTIVATION

The ability to read and infer other people's mental states can be used for competition and exploitation as well as cooperation. Mentalistic skills are the basis for Machiavellian intelligence and enable people to outwit, deceive, and manipulate others to their own advantage (Byrne & Whiten, 1989; Flinn et al., 2005). In fact, mathematical models indicate that competition promotes the evolution of sophisticated mentalistic cognition more strongly than cooperation. When cooperative interactions are more frequent than competitive ones, selection is likely to favor moderate levels of mentalizing, which are sufficient to form and maintain shared goals between cooperation partners. In a competitive social environment, however, it is crucial to have better mentalistic skills than one's opponents in order to predict and outsmart their moves (Devaine et al., 2014). This idea is consistent with evidence that competitive interactions activate mentalizing-related brain areas more strongly than cooperation (Decety et al., 2004).

To a remarkable degree, the mentalistic processes deployed in social interactions are modulated by the individual's motivational state. For example, empathic emotional responses to another person's joy or distress are favored in the context of cooperation, affiliation, and caregiving but inhibited by aggression and dominance competition. Likewise, activation of some motivational systems such as the fear and disgust systems may temporarily inhibit sophisticated mentalizing to promote rapid, stereotyped defensive responses (Liotti & Gilbert, 2011). Intriguingly, there is evidence that prosocial and hostile mentalistic abilities are negatively correlated in children, suggesting that they rely on partially distinct mechanisms (Ronald et al., 2005).

Mechanistic Cognition

Mechanistic cognition deals with objects whose behavior can be effectively predicted by applying cause–effect reasoning or input–output rules. Two mechanistic domains with a long evolutionary history are *folk physics* (the intuitive understanding of objects and physical laws) and *folk biology* (the intuitive understanding of biological organisms and their physiology; Geary, 2010). However, the same “systemizing” abilities can be employed to make sense of many other rule-based domains, from abstract ones such as mathematics and music to social ones such as law and economics (Baron-Cohen, 2003). As with mentalizing, mechanistic cognition depends on multiple mechanisms, some of which show evidence of specialized design. For instance, children process information about animals and plants differently, and do so in ways that are tailored to the characteristics of each biological category (Atran, 1998; Broesch et al., 2014; Greif et al., 2006). The domain of folk physics includes important specialized abilities such as spatial navigation, three-dimensional visualization, and tool use.

There is evidence that mentalistic and mechanistic cognition depend on different brain networks, and—even more importantly—that they tend to inhibit one another when people face social versus physical tasks (Jack et al., 2013). Indeed, there may be intrinsic tradeoffs between mentalistic and mechanistic reasoning, possibly because they employ incompatible heuristics that generate contradictory predictions; however, the exact nature of these tradeoffs is still poorly understood.

Mentalistic and Mechanistic Cognition in the Brain

Many different brain structures contribute to mentalizing; for example, the recognition of emotional expressions recruits several regions discussed previously in relation to motivation and emotion, including the ACC, DMPFC, and VLPFC (Lindquist et al., 2012). In addition, studies of mentalizing converge on a network that includes the temporoparietal junction (TPJ), superior temporal sulcus (STS), inferior parietal lobule (IPL), and inferior frontal gyrus (IFG, a subregion of the VLPFC), as well as medial prefrontal areas (Figure 2.2). These cortical structures serve a diverse set of functions—biological motion detection in the STS, imitation and action

understanding in the IPL and IFG, and so on (Frith & Frith, 2010; Schaafsma et al., 2015). The brain correlates of mentalistic cognition overlap with the *default mode network*, which includes the superior temporal lobe and TPJ, IPL, IFG, PCC and adjacent parietal regions, and medial prefrontal areas. Coordinated activation of the default network is associated with both introspective, self-referential thought and other-directed mentalizing. Mechanistic cognition is also very heterogeneous, but there is only scant evidence on the neural basis of key mechanistic skills such as folk physics and folk biology. There are indications that physical reasoning activates the *dorsal attention network*, which comprises the dorsal parietal cortex together with some prefrontal and temporal regions; its activity is weakly negatively coupled with that of the default mode network (Broyd et al., 2009; Dixon et al., 2017; Fischer et al., 2016; Jack et al., 2013).

KEY NEUROBIOLOGICAL SYSTEMS

In this section, I review a number of important neurobiological systems: brain monoamines and “social” neuropeptides, excitatory and inhibitory neurotransmitters (glutamate and GABA), sex hormones, and the stress response system. Instead of detailing these systems at the molecular level—receptor types, synthesis and degradation enzymes, specific genetic variants—I focus on their main biological functions and their role in the psychological processes discussed so far. One should keep in mind that, despite decades of animal and human research on these systems, our understanding of their role in behavior and personality is still tentative and provisional. Reasons for this state of things include the prevalence of small-sample studies (that inevitably produce many spurious and contradictory findings); the technical challenges of measuring neurobiological function in vivo; the methodological and statistical problems involved in detecting the effects of individual genes; and the remarkable complexity of the systems themselves (see Chapter 3). Acronyms for the neurotransmitters, hormones, and neuroendocrine systems discussed in the book are listed in [Box 2.2](#).

Box 2.2

Acronyms for Neuroendocrine Systems, Hormones, and Neurotransmitters

5-HT: serotonin
ACh: acetylcholine
ACTH: adrenocorticotrophic hormone
AVP: vasopressin
CRH: corticotropin-releasing hormone
DA: dopamine
DHEA: dehydroepiandrosterone
DHEAS: dehydroepiandrosterone sulfate
FSH: follicle-stimulating hormone
GABA: γ -aminobutyric acid
GnRH: gonadotropin-releasing hormone
HPA: hypothalamic-pituitary-adrenal (axis)
HPG: hypothalamic-pituitary-gonadal (axis)
HPT: hypothalamic-pituitary-thyroid (axis)
IGF-1: insulin-like growth factor 1
LH: luteinizing hormone
NE: norepinephrine
OT: oxytocin

Brain Monoamines

DOPAMINE

The main sources of dopamine (DA) in the brain are the ventral tegmental area (VTA) and the substantia nigra (SN). Neurons in the SN project to the dorsal striatum, whereas VTA neurons project to the prefrontal cortex (*mesocortical* pathway), the ventral striatum (including the NAcc), and other motivation- and emotion-related regions such as the hippocampus, cingulate cortex, and amygdala (*mesolimbic* pathway; see [Figure 2.4](#)). In turn, many of these structures project back to dopaminergic nuclei. Dopamine has been classically involved in reward learning; more broadly, it mediates psychological reactivity to both positive and negative events, likely by encoding the salience of stimuli and/or the individual's ability to control the outcome of actions. Dopaminergic activity modulates multiple motivational systems—chiefly the aggression system (by facilitating active approach), the mating system (by increasing sexual desire and facilitating orgasm), and the acquisition system (by increasing the reward value of desirable items).

Dopamine also plays a complex role in regulating hunger and food intake (Frank et al., 2009; Friston et al., 2014; Giuliano & Clément, 2005; Kayser et al., 2014; Miczek & Fish, 2006; Moore & Depue, 2016; Panksepp, 2011; Preston & Vickers, 2014; Ramos et al., 2005; Toronchuk & Ellis, 2013).

From the standpoint of self-regulation, dopamine is a key determinant of approach motivation and represents the main neurobiological substrate of the BAS. In particular, dopamine is more strongly linked to appetitive (“wanting”) than to consummatory aspects of approach motivation (“liking”). The reciprocal connections between the VTA, ventral striatum, and prefrontal cortex contribute to regulate risk sensitivity, executive flexibility, and both speed–accuracy and exploration–exploitation tradeoffs in decision-making. The effects of dopamine on these parameters may be nonlinear and depends critically on the specific site of release. In particular, prefrontal and striatal dopamine have opposing influences on some aspects of cognition and behavior; notably, prefrontal (mesocortical) dopaminergic activity is associated with behavioral persistence, whereas striatal (mesolimbic) dopamine facilitates flexibility—consistent with an executive tradeoff between flexibility and inhibition (see earlier discussion). High DA levels in the striatum—but not in prefrontal areas—have been associated with risk proneness, speed over accuracy, and exploitation over exploration; however, there is still debate concerning potential associations between *reduced* striatal dopaminergic function, impulsivity, and risk-taking (Cools, 2011; Corr et al., 2013; DeYoung, 2013; Gatzke-Kopp, 2011; Kayser et al., 2014; Sullivan & Brake, 2003; van Schouwenberg et al., 2010; Wiecki & Frank, 2013; Zisner & Beauchaine, 2015). Finally, dopaminergic function may affect mentalizing, as increased DA levels seem to reduce empathic concerns about inflicting harm to others (Crockett et al., 2015).

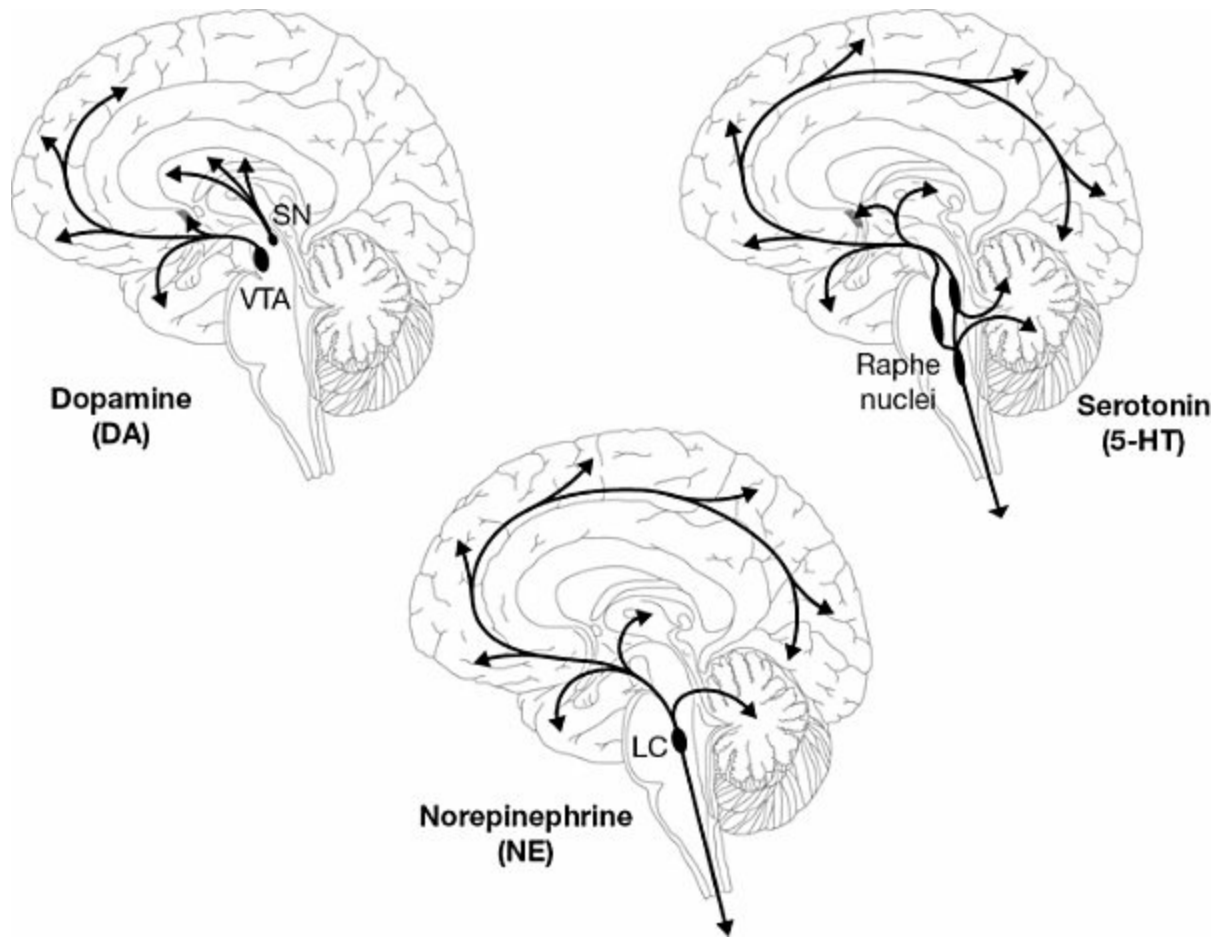


Figure 2.4. The main monoaminergic pathways in the brain. LC = locus coeruleus. SN = substantia nigra. VTA = ventral tegmental area.

SEROTONIN

In the brain, serotonin (5-HT) is synthesized in the brainstem raphe nuclei and distributed widely across cortical and subcortical structures (Figure 2.4). This monoamine is derived from tryptophan, an amino acid that cannot be synthesized by humans and must be acquired from food. The original biological function of serotonin is that of regulating metabolism and energy allocation (Andrews et al., 2015); as a neuromodulator, it shows a broad range of effects, including a complex pattern of antagonistic and synergistic interactions with dopamine. For example, serotonin suppresses dopaminergic activity in the VTA and SN and inhibits dopamine release in the ventral striatum, but facilitates reward processing in the dopaminergic system (Cools et al., 2011; Rogers, 2011). Serotonin is involved in aversive learning and threat avoidance and modulates several motivational systems—including key defensive systems such as fear, security, and disgust. Specifically,

serotonergic activity potentiates disgust and upregulates security motivation by promoting anticipatory anxiety; at the same time, it downregulates the expression of some fear and panic responses and suppresses reactive aggression, but not proactive aggression. In the mating domain, serotonergic activity inhibits sexual desire, arousal, and orgasm; serotonin also reduces hunger and suppresses food intake (Canteras & Graeff, 2014; Corr et al., 2013; Giuliano & Clément, 2005; Graeff, 2004; Lam et al., 2010; Manuck et al., 2006; Mickzek & Fish, 2006; Pfaus, 2009; Ramos et al., 2005; Toronchuk & Ellis, 2013; van Goozen et al., 2007; Vicario et al., 2017).

The serotonergic system plays important roles in self-regulation, mainly as a source of behavioral constraint and withdrawal motivation. High serotonin levels are robustly associated with inhibition in response to threats and aversive stimuli (i.e., BIS activation), risk aversion, and reduced time discounting with increased preference for delayed rewards (Carver et al., 2015; Cools et al., 2011; Corr et al., 2013; Moore & Depue, 2016; Rogers, 2011; Tops et al., 2010). In contrast, the role of serotonin in executive flexibility is unclear, and the empirical findings are inconsistent (Cools et al., 2011; Faulkner & Deakin, 2014). Serotonin has also been found to promote prosocial and cooperative behavior, responsivity to social norms, and empathic aversion to inflicting harm to others (Bilderbeck et al., 2014; Crockett et al., 2015; Rogers, 2011).

NOREPINEPHRINE

Norepinephrine (NE) is released in the general circulation by the medulla of the adrenal gland and plays a central role in the stress response system. In the central nervous system, norepinephrine is produced in the locus coeruleus (LC) and transported to many brain areas, with reciprocal connections between the LC and prefrontal and cingulate regions, amygdala, hypothalamus, and other structures (Figure 2.4). Norepinephrine increases arousal and vigilance and facilitates reactive aggression (Miczek & Fish, 2006; Moore & Depue, 2016). Phasic NE spikes are associated with novel or surprising stimuli and contribute to reorient attention and behavior toward the unexpected event. They also promote engagement and performance by amplifying task-related information while filtering out irrelevant stimuli. In contrast, high tonic levels of the neurotransmitter depress phasic activity and facilitate exploration, flexibility, and distractibility. The balance of exploration versus exploitation in noradrenergic activity is regulated by inputs from other regions of the brain, primarily the ACC and OFC (Adams

et al., 2012; Aston-Jones & Cohen, 2005; Bouret & Sara, 2005; Yu & Dayan, 2005).

Neuropeptides

OXYTOCIN

Oxytocin (OT) is produced in the supraoptic and paraventricular nuclei of the hypothalamus (SON and PVN). From there, oxytocin is released as a hormone in the bloodstream through the pituitary gland and as a neuromodulator in various brain regions including the striatum, amygdala, hippocampus, and VTA. The overarching psychological function of oxytocin is to motivate and maintain social bonds, from mother–child relationships to cooperation in extended groups. Oxytocinergic activity modulates the attachment and caregiving systems, the mating system (e.g., OT is released during orgasm), the pair bonding systems, the affiliation system, and the reciprocity system. Importantly, oxytocin is released in response to both relational opportunities (e.g., a beginning romantic relationship) and relational threats (e.g., abandonment, rejection), and seems to specifically promote investment in developing or vulnerable social bonds (Bos et al., 2012; Brown et al., 2012; Crespi, 2016a; Gangestad, 2016; van Anders et al., 2011). High OT levels also tend to suppress hunger and food intake (Lam et al., 2010; Lawson et al., 2015). Oxytocin works as an enhancer of mentalizing—it facilitates empathy, emotion reading, and social learning. Depending on context, oxytocin can enhance trust and cooperation with social partners; at the same time, it has been found to inhibit cooperation with strangers and out-group members (Bos et al., 2012; Crespi, 2016a; Hurlemann et al., 2010).

VASOPRESSIN

The peptide vasopressin (AVP) is closely related to oxytocin; like oxytocin, it is synthesized in the SON and PVN, released in the pituitary, and transported throughout the brain as a neuromodulator. However, the two peptides have largely contrasting effects, for example on patterns of neural activity in the amygdala. Vasopressin is involved in social aggression and promotes mating, courtship behaviors, and status competition (especially in males). There is some evidence that AVP also contributes to caregiving and may specifically regulate aggression in the context of offspring defense. Recent evidence in

humans suggests a potential involvement of AVP in high-risk cooperation among men (Bos et al., 2012; Brunnlieb et al., 2016; Panksepp, 2011; Sanchez et al., 2009; Toronchuk & Ellis, 2013).

OPIOIDS

Several opioid molecules are produced in the brain, mostly in the hypothalamus and brainstem. Opioid receptors are highly expressed in the brainstem, amygdala, striatum, and prefrontal cortex (especially the OFC). β -endorphin is the most potent and well-studied of the endogenous opioids and contributes to modulate pain (with an analgesic function) as well as reward processing (particularly the consummatory or “liking” component). Opioids are crucially involved in the formation and maintenance of social bonds. Grooming, affectionate touch, and laughter all stimulate opioid release; more broadly, opioid levels increase in response to closeness and affiliative behaviors and drop following separations. Low opioid levels increase the need for social stimulation and promote activation of the attachment, caregiving, and affiliation systems. Opioids also participate in play and mating (sexual pleasure) and have stimulatory effects on hunger and food intake (Machin & Dunbar, 2011; Moore & Depue, 2016; Panksepp, 1998, 2011; Yeomans & Gray, 2002).

Glutamate and GABA

GLUTAMATE

The amino acid glutamate is the main excitatory neurotransmitter and one of the most abundant molecules in the brain. Glutamate is involved in virtually all brain processes and plays a key role in regulating neural plasticity. As a neuromodulator, glutamate tends to increase the excitability of neural networks; glutamatergic projections to the NAcc from prefrontal and limbic areas contribute to regulate approach and avoidance motivation, with important effects on the BIS, BAS, fear, and aggression systems (Floresco, 2015; Miczek & Fish, 2006; Panksepp, 2011). Both dopaminergic and serotonergic neurons have been found to co-release glutamate, which may contribute to modulate the psychological effects of these monoamines (Fischer et al., 2015).

GABA

GABA is produced by a widely distributed system of interneurons in the brain cortex and is the main inhibitory neurotransmitter/neuromodulator. GABA has broad anxiolytic effects and plays a crucial role in dampening the activity of defensive systems such as the BIS and the security system (Moore & Depue, 2016). GABAergic activity is also involved in the modulation of the aggression system; the relation between GABA and aggression seems to be nonlinear, with enhanced aggression at moderate GABA levels but suppressed aggression at high levels (Miczek & Fish, 2006).

The Stress Response System

The stress response system is an integrated neuroendocrine network that coordinates the organism's response to a wide range of metabolic and behavioral challenges. Its main components are the autonomic nervous system and the hypothalamic-pituitary-adrenal (HPA) axis. The whole-body adjustments mediated by the stress response system are captured by the term *allostasis*—the maintenance of stability through change (Sterling & Eyer, 1988). The system receives input from the amygdala, hippocampus, and prefrontal and cingulate regions, but also from other organs and systems in the body (including the immune system). It responds to both threats and opportunities by reorienting attentional focus, increasing the organism's readiness for action (e.g., by regulating heart/respiratory rate and blood flow to various organs), mobilizing energy by triggering glucose release into the bloodstream, shifting the balance between different memory and learning processes, suppressing or enhancing reproductive physiology, regulating immune function, and many other effects (see Del Giudice et al., 2011; Ellis & Del Giudice, 2014; Flinn, 2006, Flinn et al., 2011; Gunnar & Quevedo, 2007; Miller G. E. et al., 2011). In short, the stress response system functions as a central node in the integration of physiology and behavior and shows extensive bidirectional interactions with all the neurobiological systems described in this section.

AUTONOMIC NERVOUS SYSTEM

The autonomic nervous system is composed of two branches, *sympathetic* and *parasympathetic*. Autonomic branches innervate most bodily organs; their complementary physiological effects result in a dynamic equilibrium in which parasympathetic activation suppresses sympathetic-induced arousal.

The parasympathetic system employs acetylcholine (ACh) as its main neurotransmitter. Parasympathetic activity promotes digestive and reproductive functions, reduces the heart and respiratory rate, and sustains attention and social engagement (Lovallo & Sollers, 2007; Porges, 2001, 2007). Disengagement of the parasympathetic “brake” provides a way to quickly increase arousal and reorient attention toward unexpected events. Parasympathetic activation is involved in the regulation of many social behaviors including mating, caregiving, affiliation, and play; it is also a robust physiological correlate of disgust (Ottaviani et al., 2013). Signaling in the sympathetic system is based on two catecholamines, epinephrine (E) and NE. Sympathetic activity is coordinated by the LC (see earlier discussion); from there, its activation follows two routes—a fast pathway mediated by direct innervation of the brain and other organs, and a slower pathway in which sympathetic fibers project to the medulla of the adrenal gland, which in turn releases E and NE in the circulatory system. Sympathetic effects on physiology include increased heart and respiration rate, increased blood supply to skeletal muscles, glucose release in the bloodstream, and suppression of vegetative functions.

Motivational systems often recruit both branches of the autonomic system with complementary functions. For example, parasympathetic activation promotes sexual arousal, whereas sympathetic activity peaks during orgasm (Giuliano & Clément, 2005; Krüger et al., 2003; Salonia et al., 2010). In responding to danger and threats, the sympathetic branch promotes fight-or-flight behaviors, whereas the parasympathetic one mediates freezing and fainting (Porges, 2007). While the sympathetic and parasympathetic systems usually operate in a reciprocal fashion, they can also become strongly activated at the same time. This occurs during attentive immobility (freezing) and in situations that demand high arousal and readiness to action but also tight self-control (e.g., skydiving; Allison et al., 2012).

HPA AXIS

The HPA axis mounts long-term responses to metabolic and social challenges through the release of cortisol, a hormone with a wide range of physiological and behavioral effects. Hypothalamic neurons secrete corticotropin-releasing hormone (CRH) and AVP in response to inputs from multiple brain regions such as the LC, amygdala, and hippocampus. CRH and AVP are transported to various brain areas—including the LC, where they modulate sympathetic activity—and to the pituitary gland, where they stimulate release of

adrenocorticotrophic hormone (ACTH) in the bloodstream, together with other molecules such as β -endorphin. ACTH then reaches the cortex of the adrenal gland, where it triggers cortisol release (Figure 2.5). Cortisol begins to rise within minutes from the onset of a stressor, and some of its effects may last for several hours. Rising cortisol levels activate the glucocorticoid receptor (GR) at various levels of the axis (Figure 2.5), providing negative feedback that ultimately deactivates the system. Initially, cortisol potentiates sympathetic activity, but later contributes to dampen it and suppress some of its effects (Del Giudice et al., 2011; Flinn, 2006; Gunnar & Quevedo, 2007).

As a neuromodulator, CRH has a range of psychological effects that include facilitating anxiety, enhancing the formation of emotional memories, and suppressing hunger. Cortisol mobilizes energetic (glucose) and attentional resources, and modulates motivation and self-regulation on a broad scale (Ellis & Del Giudice, 2014; Moore & Depue, 2016). The HPA axis is activated by both threats and opportunities, such as the presence of a potential sexual partner (e.g., Roney et al., 2007). Events that involve social judgment and/or uncontrollability arouse strong, sustained HPA responses (Dickerson & Kemeny, 2004). While cortisol has important benefits, maintaining high levels of this hormone also has substantial costs for the organism; thus, periods of chronic HPA activation are usually followed by a “rebound” below the initial level that can last weeks or even months after the end of the stressor (Miller G. E. et al., 2007).

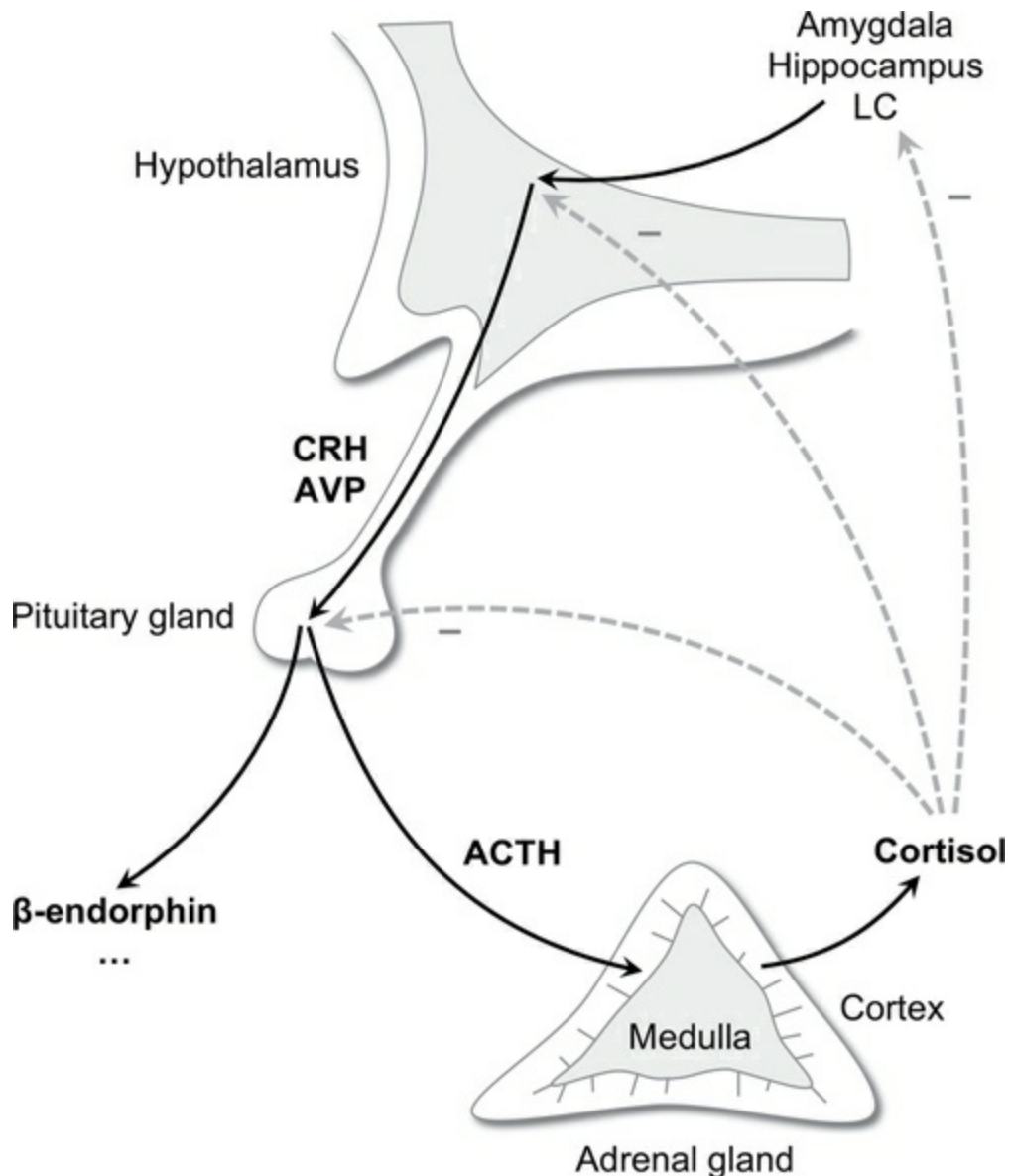


Figure 2.5. The hypothalamic-pituitary-adrenal (HPA) axis. Cortisol exerts negative feedback on the axis at multiple levels through the glucocorticoid receptor. ACTH = adrenocorticotropic hormone. AVP = vasopressin. CRH = corticotropin releasing hormone. LC = locus coeruleus.

Sex Hormones

Sex hormones are the end products of another neuroendocrine system, the hypothalamic-pituitary-gonadal (HPG) axis. Hypothalamic neurons release gonadotropin-releasing hormone (GnRH) in the pituitary gland, where luteinizing hormone (LH) and follicle-stimulating hormone (FSH) are

secreted into the blood. In response to LH/FSH signals, the ovaries and testes release androgens (mainly testosterone), estrogens (mainly estradiol), and progesterone. Androgens and progesterone are also produced in the cortex of the adrenal gland. The metabolism and action of sex hormones is complex; for example, adrenal androgens dehydroepiandrosterone (DHEA) and dehydroepiandrosterone sulfate (DHEAS) can be converted to testosterone, which in turn can be converted to estradiol. Sex hormones regulate sexual development, reproduction, and fecundity and have a striking range of behavioral and cognitive effects. There are strong bidirectional influences between the HPG and HPA axis, as well as among sex hormones, monoamines, and neuropeptides (Bos et al., 2012; Toronchuk & Ellis, 2013; van Anders et al., 2011).

The main motivational effects of testosterone are on the mating, status, aggression, and fear systems. In both sexes, testosterone promotes competition for dominance and status, mates, and resources. Specifically, testosterone stimulates aggression, suppresses fear, and is crucially involved in courtship and sexual behavior (Bos et al., 2012; Cashdan, 1995; Eisenegger et al., 2011; Harris et al., 1996; Mazur & Booth, 1998; Zilioli & Bird, 2017). Testosterone also increases approach motivation and risk-taking, particularly in competitive contexts; at the same time, it decreases trust and affiliation (Apicella et al., 2015; Bos et al., 2012; Crespi, 2016a). Testosterone levels in men decline when they enter committed romantic relationships (e.g., marriage) and become fathers; this is consistent with the idea that testosterone plays a role in courtship and relationship initiation but interferes with long-term pair bonding and parental investment. There is some evidence that similar effects may occur in women as well (Barrett et al., 2013; Carré & Archer, 2018; Gettler, 2014; Roney & Gettler, 2015). In women, estrogen and progesterone levels change across the menstrual cycle and jointly contribute to regulate motivational states. Estrogens and progesterone often have opposite effects on psychological processes; for example, estrogens tend to increase sexual motivation and decrease hunger and food intake, whereas progesterone inhibits desire and promotes hunger (Hirschberg, 2012; Roney & Simmons, 2013).

Sex hormones and social neuropeptides are functionally linked: estrogens increase oxytocin synthesis and release, while testosterone has a stimulating effect on vasopressin. Indeed, neuropeptides probably mediate some of the psychological effects of sex hormones. In general terms, activity of the

estrogen–OT pathway tends to promote nurturance, prosocial affiliation, and mentalizing, whereas the testosterone–AVP pathway promotes self-oriented behaviors, reduced mentalizing, and a focus on threats and rewards (Bos et al., 2012; Crespi, 2016*a*; van Anders et al., 2011). However, things are likely more complicated than suggested by this schematic account; for example, testosterone and vasopressin can actually promote caregiving motivation, especially in situations that involve protection of a vulnerable partner or child (Gettler, 2014; van Anders et al., 2011). There is also initial evidence that testosterone can stimulate oxytocin release and vice versa, at least in caregiving contexts (Gangestad, 2016; Weisman et al., 2014).