Association for Psychological Science

Perspectives on Psychological Science
8(2) 119–141

DOI: 10.1177/1745691612474317
http://pps.sagepub.com
SAGE

Navigating Into the Future or Driven by the Past

Martin E. P. Seligman[1], Peter Railton[2], Roy F. Baumeister[3], and Chandra Sripada[2]
[1]University of Pennsylvania, Philadelphia; [2]University of Michigan, Ann Arbor; and [3]Florida State University, Tallahassee

Abstract

Prospection (Gilbert & Wilson, 2007), the representation of possible futures, is a ubiquitous feature of the human mind. Much psychological theory and practice, in contrast, has understood human action as determined by the past and viewed any such teleology (selection of action in light of goals) as a violation of natural law because the future cannot act on the present. Prospection involves no backward causation; rather, it is guidance not by the future itself but by present, evaluative representations of possible future states. These representations can be understood minimally as "If *X*, then *Y*" conditionals, and the process of prospection can be understood as the generation and evaluation of these conditionals. We review the history of the attempt to cast teleology out of science, culminating in the failures of behaviorism and psychoanalysis to account adequately for action without teleology. A wide range of evidence suggests that prospection is a central organizing feature of perception, cognition, affect, memory, motivation, and action. The authors speculate that prospection casts new light on why subjectivity is part of consciousness, what is "free" and "willing" in "free will," and on mental disorders and their treatment. Viewing behavior as driven by the past was a powerful framework that helped create scientific psychology, but accumulating evidence in a wide range of areas of research suggests a shift in framework, in which navigation into the future is seen as a core organizing principle of animal and human behavior.

Keywords

prospection, telos, default circuit, free will, consciousness, psychotherapy

Much of the history of psychology has been dominated by a framework in which people and animals are driven by the past. In this picture, past history, present circumstance, and inner states drive behavior, much as in a classical dynamical system the vector sum of forces operating on and within a particle uniquely determines its trajectory. We suggest an alternate framework in which people and intelligent animals draw on experience to update a branching array of evaluative prospects that fan out before them. Action is then selected in light of their needs and goals. The past is not a force that drives them but a resource from which they selectively extract information about the prospects they face. These prospects can include not only possibilities that have occurred before but also possibilities that have never occurred—and these new possibilities often play a decisive role in the selection of action.

This is not remotely a novel idea. It is a feature of common sense, and in a number of areas of contemporary psychology this idea now plays an important role, at least implicitly. Our goal is to make this implicit notion as explicit as possible and to use it as a framework for integrating many lines of research in contemporary psychology—including, learning, cognition and memory, emotion and motivation, and self-control and decision making.

We have organized our argument as follows: The first section is about the history of casting teleology out of science and how this culminated in behaviorism's and Freudian psychology's common premise that action is driven by the past. This allows us to review the threads of evidence that led away from this premise, as experimental evidence of purposive action and of the limited effects of the past kept cropping up. In the second section, we adduce a priori and empirical grounds for thinking that prospection would be efficient, effective, and favored in natural selection. Here we point to a powerful convergence between ideal models of learning and decision making and actual psychological processes. This convergence contrasts, however, with the recent emphasis on the deficiencies of prospection. In the third section, we consider emerging

Corresponding Author:
Martin E. P. Seligman, University of Pennsylvania, 3701 Market Street, Philadelphia, PA 19104
E-mail: seligman@psych.upenn.edu

evidence of prospection as a core brain process with identifiable networks of brain circuitry. In the last section, we speculate about issues that are opaque to the driven-by-the-past framework but are transparent to the navigating-the-future framework: subjectivity and consciousness, freedom of the will, and mental disorders involving failures of prospection.

The Fall and Rise of Teleology

Thinking is designed for doing, as William James famously asserted (James, 1890, 1:33), and because doing affects the future, never the past, such thinking will be designed for what Gilbert (2006), Gilbert and Wilson (2007), and Buckner and Carroll (2007) called “prospection,” the mental simulation of future possibilities. We call such accounts “teleological,” meaning explanation by selection in light of values and goals, where “telos” means “end.” A good prospector must know more than the physical landscape—what is to be found where, with what probability—but also at what cost in effort and risk and with what possible gain. The prospecting organism must construct an *evaluative landscape* of possible acts and outcomes. The organism then acts through this evaluative representation, electing action in light of these prospects. And the success or failure of an act in living up to its prospect will lead not simply to satisfaction or frustration but to maintaining or revising the evaluative representation that will guide the next act. To be sure, learning and memory necessarily reflect past experience. But at any given moment, an organism’s ability to improve its chances for survival and reproduction lies in the future, not the past. So learning and memory, too, should be designed for action. These capacities actively orient the organism toward what might lie ahead and what information is most vital for estimating this. Impinging sensation must be registered selectively in ways that enhance information value and permit flexibility and improvisation to meet new threats and opportunities—not simply serving to entrench and trigger habits.

As intuitive as such “teleological” accounts of thinking and acting might seem, they are decidedly not the accounts that prevailed throughout much of the history of modern psychology. Instead, teleology has largely been anathema in the field, to be replaced wherever possible with mechanical determination by the past. Why?

To understand this requires going back to the modern origins of experimental psychology in 16th- and 17th-century philosophy and science, a time when the project of banishing teleology from the study of nature was indeed indispensable for scientific progress. Aristotelian and neo-Aristotelian explanations invoked so-called final causes—heavy bodies fell to earth because the earth is their natural resting place, and the heavier they were, the stronger their affinity and the faster they fell. Such an explanation appeared to fit everyday observation and provided a satisfying way of answering “Why?”

However, once Galileo and other early modern experimentalists and astronomers learned how to measure nature with quantitative accuracy and use mathematics to formulate law-like regularities, they found such teleological accounts at odds with observable facts and explanatorily uninformative. Bacon wrote, “Inquiry into final causes is sterile, and, like a virgin consecrated to God, produces nothing” (1623/1959, p. 3.5). The English philosopher Hobbes, rightly impressed by Galileo’s discoveries, sought to apply a Galilean method to create a human science (1651/1994). Rather than explain political order in terms of an Aristotelian “natural order,” he broke society into its component parts, individuals, and used their “laws of motion”—essentially, appetites and aversions—to explain the emergence of social combination. Hobbes’s method offended orthodoxy, but it had the advantage of seeming obvious to the scientifically minded and set the tone for much of what was to come. As the great French mathematician Pierre-Simon Laplace later argued, “We may regard the present state of the universe as the effect of its past and the cause of its future”—more strictly, given the positions and momenta of all the individual particles in the universe at a particular moment and the laws of motion, an intellect vast enough could foresee the entire future of the universe (1814/1951, p. 4). No divination of nature’s ostensible aims, no teleology, was required to predict or explain the totality of its behavior.

In the 19th century, Darwin, Marx, Freud, and James each made a contribution to ridding his domain of teleological vestiges. Darwin (1859/1956) replaced the apparently natural design of organisms with chance variation and selective retention. Marx (1844/1988) replaced Hegelian teleology with an account of historical change as driven by technological innovation and social contest. Freud (1901/2002, 1920/1975) replaced Aristotle’s rational soul aiming at the good with a complex psychodynamic of competing and largely unconscious drives—seemingly inexplicable dysfunctional or accidental behavior was caused by unresolved past conflicts, often originating in childhood, playing themselves out below the level of awareness.

And in James’s complex portrayal of psychic processes, he argued that no finalistic valuation or “clairvoyant or prophetic power” need be found to account for the seeming “end-oriented” behavior of humans and animals (1890, 2:384). Indeed, in this respect, humans are simply scaled-up animals: “Not one man in a billion, when taking his dinner, ever thinks of utility”; instead, “Man has a far greater variety of impulses than any lower animal,” yet “any one of these impulses, taken in itself, is as ‘blind’ as the lowest instinct can be” (1890, 2:390). Association explains the appearance of future-directedness: “*every instinctive act, in an animal with a memory, must cease to be ‘blind’ after being once repeated*,” and what appeared to be “foresight of its ‘end’” was simply the triggering of this associated memory by the circumstances leading to the act itself (1890, 2:390).

These theories were major achievements, and their explanation of how objectionable elements of teleology could be purged from the study of biological and human phenomena contributed greatly to the advance of knowledge. But none would have advocated expunging purpose, or mentalism, in their entirety.

The rise and fall of behaviorism

That project was left for behaviorism. Its chief motivation was methodological: to permit the construction of a yet more scientifically rigorous and quantitative theory of behavior. Among other innovations, it embraced the use of "stimulus" and "response" as the scientific terms for understanding action. These terms themselves, unlike more neutral terms such as "event" and "action," displayed the driven-by-the-past framework. The great experimentalist Karl Lashley wrote, "The problem which confronts the behaviorist is to find in the physical world deterministic relations between non-qualitative [i.e., nonmentalistic] discrete entities in time and space" (Lashley, 1923, p. 329) to replace the unquantifiable elements of introspective psychology with observable stimuli mechanically causing observable responses. He saw as "the most serious defect in current psychology" the fact that "current psychological language is a weird composite of teleological and mechanistic terms," even though the "two systems, mechanistic explanation and finalistic valuation, stand out as incompatible points of view" (1923, pp. 346, 349). For Lashley at the time, and for many behaviorists on into the post–World War II period, "finalistic valuation" or "teleology" was incoherent—a kind of causation of the present by the future. A properly scientific account of behavior "would not reveal an influence of the future on the present, nor does the behaviorist account" (1923, p. 349). As Clark Hull wrote in 1943, Jamesian pragmatism, Watsonian behaviorism, Russian "reflexology," and quantitative techniques were uniting "to produce in America a behavioral discipline which will be a full-blown natural science" (Hull, 1943b, p. 273).

The methodological centrality of constructing strictly quantitative laws perhaps explains why the discoveries of quantum mechanics, or the much earlier recognition in physics of the intractability of calculating trajectories in a classical system once it involves three or more interacting bodies (see Wolfram, 2002), did not deter behaviorists from following determinism as a "canonical conviction" (Bargh & Ferguson, 2000, p. 925). While the rest of science moved toward the view that prediction, and perhaps explanation as well, can at best be probabilistic, for the behaviorist, the scientific route was to push the Laplacian model as far as possible.

To be sure, there are many variations in the details of different learning theories, but perhaps our characterization captures the essentials. In principle, it seemed, something like this picture must be right. No objectionable reference to teleological evaluation should be needed to explain the rat's turning left in the maze, despite the apparently forward-looking intelligence of its behavior. Behaviorists assumed that human behavior could likewise be explained without teleology. Human drives and reinforcers might be more varied, and humans might be capable of longer stimulus–response chains and wider stimulus generalization, but there were to be no exceptions to the mechanical model—notions such as expectations of as yet nonexistent future events were an invitation to untestability at best, obscurantism and incoherence at worst.

Many are now tempted to say that the failure of behaviorism lay in its overreach, trying to use a theory that worked for rats and pigeons in the experimental setting to explain human psychology in unconstrained situations (Bargh & Ferguson, 2000; B. Schwartz, Schuldenfrei, & Lacey, 1978). But we believe that the crucial failure was in eschewing teleological explanation, which followed directly from the exclusion of mental events in favor of drives and habits. This failure is perhaps clearest on its home ground: Behaviorist learning theory did not even work for white rats in the laboratory.

How did this seemingly self-evident approach fail? From early behavioral research on, white rats, behaviorism's subjects of convenience, looked suspiciously future oriented in ways that did not seem easily reducible to past habits—even in the venerable T-maze. Rats who had been reinforced for making the response of turning left to get food had the right motor cortex of their brains ablated, making it physically impossible to turn left. If reinforcement were about motor habits, they should have walked up to the choice point and stalled—unable to "emit" the instrumental response. Instead, they walked up to the choice point and promptly made 270° right turns (Lashley, 1929), even if this meant rolling, somersaulting, or dragging paralyzed limbs. These experiments, and incidental reports of chance observations, for example, of maze-trained rats escaping the starting box and walking diagonally across the top of the maze, directly to the food (Lashley, 1929), pointed clearly away from the idea that behavior was under the control of past motor "habits," suggesting instead that an acquired "cognitive map" governed navigation flexibly, permitting goal-directed behaviors of unprecedented kinds.

This idea gained further support from doing what behaviorists most strenuously urged: making close observations of actual behavior. Psychologists who looked up from their automatic recording systems noticed that, when not "overtrained," rats at the choice point would occasionally turn their heads to the left, to the right, and back again. This observation, dubbed "vicarious trial and error," along with a suite of experimental evidence, suggested that "intervening brain processes are more complicated, more patterned, and often, pragmatically speaking, more autonomous than [is accepted by] the stimulus-response psychologists" (Tolman, 1948, p. 192).The furry "billiard ball" indeed has a mind of its own.

Even more trouble for banishing what had been seen as teleology came from Pavlovian conditioning. Pavlovian conditioning was seen by behaviorists as the nonmentalistic operationalization of pure associationism. Any teleological or mentalistic suggestion was removed by calling what was conditioned a "reflex." It was supposed to occur based on the mere repeated temporal contiguity of the conditioned stimulus (CS) and the unconditioned stimulus (UCS) that elicited an unconditioned reflex. Pavlov's original Russian was mistranslated in such a way as to underscore the lack of teleology. His "conditional" and "unconditional" were rendered "conditioned" and "unconditioned" (McGuigan & Ban, 1987, p. 138), and this shift to the past participle brought with it the connotation of a fixed relation, whereas "conditional," Pavlov's term, allows the "if–then" representation of possibilities.

Translation might seem beside the point—all that counts are the experiments. But in this case, words matter—as experiments later showed. It turned out that informationally redundant but nevertheless contiguous "conditioned stimuli" did not "condition" (the mistranslation had morphed into a verb). Even though a tone (CS) often occurred paired with a shock (UCS), if the base rate of shock in the absence of the tone was as high as the probability of the shock in the presence of the tone, no conditioning took place (Rescorla, 1968). Degree of conditioning was found to be an exquisitely sensitive mirror of how much information the CS gave about the likelihood of the UCS (Rescorla, 1988). This suggests not the passive ingraining of a blunt associative "connection" but the dynamic learning of a fine-grained *conditional probability*–exactly what one would expect if animal learning operated via an "if CS, then UCS" expectation.

The final fight over teleology in learning theory, underappreciated at the time, came down to avoidance learning and two-process theory. Avoidance learning, to the uninitiated, looks exactly like learning a forward-looking if–then conditional: "if I jump, I will avoid future shock." Animals are first given escape training in which a tone is followed 5 seconds later by a shock. The animals can escape the shock by jumping over a barrier to the safe side. Jumping over the barrier also turned off the signaling tone. Animals soon learned to jump as soon as the tone went on and before the shock would have gone on. This was called *avoidance learning* because the animals avoided getting any more shock by jumping as soon as the tone went on—which both prevented shock and terminated the tone. If any animal behavior looks future oriented, it is avoidance, so reducing it to the present and the past was considered a tour de force for learning theory (Rescorla & Solomon, 1967).

Here is how the reduction worked. The behaviorists denied that the animals expected anything at all, claiming that a nonoccurrent event—the avoided shock—could have no role in conditioning. Instead, a present event, the tone coming on, had become fear evoking by the usual process of Pavlovian conditioning of pairing with shock. So jumping was actually reinforced by getting the fearful tone to stop and was not at all motivated by avoiding a shock that never came.

By contrast, the cognitive theorists claimed that the animals learned that the tone predicted shock and thus acquired an *expectation* that if they jumped they would get no shock (Seligman & Johnston, 1973). That expectation—an if–then conditional about a merely possible outcome—was in itself sufficient to guide the jumping behavior.

The stage was set. This was a rare instance of a head-on collision between two theories that allowed a crucial test, in this case by extinction. Once the animals became steady jumpers, they never got shocked again, and so tone was no longer paired with shock. This is called "Pavlovian extinction," a phenomenon that should cause the tone to lose its fearful properties. If the cause of jumping were to turn off the fearful tone, as the behaviorists contended, the animals should stop jumping once they had numerous trials in which fear of the tone was extinguished. If, conversely, they were jumping to prevent the anticipated shock, as the cognitive theorists contended, the animals should keep jumping. Their if–then expectation that jumping would prevent shock was borne out every time they jumped—extinction breaks the *association* of tone and shock, but is not *disconfirmation* of the conditional. So if the conditional guides their behavior, the jumping behavior should persist.

This is exactly what happened. Hundreds of trials later, the animals were still jumping. The simplest explanation—to which no adequate counterresponse by learning theorists was offered—is that the animals had acquired a well-confirmed evaluative representation of the future (cf. Denrell, 2007, for risk-averse sampling accounts; and Erev & Barron, 2005, for a cognitive strategies account).

Behaviorist learning theory gradually gave up its ambition of providing "a set of theories to explain all behavior" (Rescorla, 1988, p. 158). Wholesale antimentalism was abandoned, and "expectations" were introduced to explain conditioned behavior (Colwill & Rescorla, 1985). The classical picture of conditioning itself underwent increasingly drastic revision, eventually resulting in the view that contiguity and repetition play no special role, that features of the UCS need not be transferred to the CS, and that "the organism is better seen as an information seeker using logical and perceptual relations among stimuli, along with its own preconceptions, to form a sophisticated representation of the world" (Rescorla, 1988, p. 154). Rats, it seems, are more predictable when we postulate that they act through complex expectation-based representations of possible actions and outcomes, behaving more like inquirers actively seeking to anticipate the future than creatures of habit-channeled drives. Perhaps humans should be given as much credit?

Thus learning theory itself, we have argued, led eventually to the conclusion that animal behavior could not be explained without positing "forbidden" internal representations of nonactual futures. But why were they "forbidden"? A conceptual error seems to have animated behaviorism, in which something genuinely suspect—a metaphysical teleology of causation backward in time, of the present by the future—was conflated with something not at all mysterious, namely, guidance by a system bearing causal and evaluative information about possible futures.

Cognitivism

Some experimental psychologists recognized the reality of expectations much earlier and sought to show the scientific respectability and importance for learning theory of expectations (Brunswik, 1951; Postman, 1951; Rotter, Fitzgerald, & Joyce, 1954; Tolman, 1948). And among those psychologists primarily interested in human behavior, expectancies were often accepted as important with less controversy. The study of self-fulfilling prophecies and expectancy effects, for example, emphasized that even wrongheaded assumptions about the future can exert a causal influence on behavior, thereby in some cases helping themselves to come true (Merton, 1949; Rosenthal & Jacobson, 1968). Bandura's (1977) theory of self-efficacy, which introduced a cognitive and agentic self into the behaviorist account, rested on the interplay between two types of expectancies. In particular, he proposed that self-efficacy involves being confident that one can perform the requisite behavior (efficacy expectation) and that the action will produce the desired outcome (outcome expectation). Mischel's (1968, 1973) effort to switch personality theory away from emphasizing stable, broad inner traits to resting instead on situational if–then strategies likewise assigned a prominent, indispensable place to expectancies. And Beck's pioneering development of cognitive therapy focused centrally on understanding how the distorted expectations of those experiencing depression or other psychological disorders contributed to their condition (Beck, Rush, Shaw, & Emery, 1979). Changing these expectations through cognitive therapy, including imaginative simulation of possible futures (Beck, Freeman, & Davis, 2003), is an important example of how, in our view, greater attention to prospection as a central organizing element can have therapeutic benefits, a point to which we will return in the final section.

But although all cognitivists by definition rejected the antimentalism of behaviorism, a number of contemporary cognitive scientists, especially those interested in "automaticity," still see important continuities with the driven-by-the-past premise of behaviorism. Lest the reader think we are attacking straw men, consider the following:

> the [contemporary] social-cognitive approach to higher mental processes, like cognitive science in general, shares with behaviorism a basic deterministic stance toward psychological phenomena. By determinism we mean, quite simply, the position that for every psychological effect (e.g., behavior, emotion, judgment, memory, perception) there exists a set of causes, or antecedent conditions, that uniquely lead to that effect. . . . Although [the] distinction between the two schools is certainly substantial and consequential, behaviorists and cognitive scientists do share certain assumptions about the nature of human volition and educe them from the same general philosophical foundations. (Bargh & Ferguson, 2000, p. 925)

Freudianism and its discontents

As the conceptual and explanatory space allowed by behaviorism was enlarged, so had Freudian psychology evolved into more eclectic versions of clinical practice. In this case, too, there was movement beyond the orthodox framework of drives. Analytic theorists felt free to introduce novel concepts, models of development and personality, and therapeutic techniques. However, what often remained in theory and practice was the attempt to explain present thought and action as the interplay of unresolved and largely unconscious conflicts, grounded in the distant past.

The psychoanalytic assumption that such unresolved conflicts drive adult personality is much less amenable to testing than the simpler mechanisms posited by the behaviorists. And analytic psychotherapy itself is a lengthy and multifaceted process, impossible to standardize. Even so, and even though recent experimental research has provided increasingly strong and detailed information about the importance of unconscious processes (Hassin, Uleman, & Bargh, 2003; Shevrin, Bond, Brakel, Hendel, & Williams, 1996), 100 years of psychoanalytic practice aimed at uncovering repressed childhood conflicts has failed to provide convincing evidence of efficacy.

More telling, perhaps, is the accumulating weight of carefully done longitudinal studies that have found disappointingly small effects of childhood events on a range of adult behaviors. For example, in an 8,000 twin-pair study of the onset of adult depression, childhood events had almost no predictive value, whereas genetic factors had great predictive power and recent events had moderate predictive power (Kendler, Walters, & Kessler, 1997). What the careful reporter of the psychoanalytic literature would be entitled to conclude is that remote past experiences, even when they seemingly are important, greatly underdetermine present feeling or action and may not point to a particular underlying psychodynamic that must be unearthed if dysfunctional behavior is to be changed. Instead, it seems, fetal and early childhood exposure to bad events, for example, has predictive power precisely because it appears to affect the way in which novel information is processed and projected (Knudsen, Heckman, Cameron, & Shonkoff, 2006).

The failure of behaviorism's favored mechanisms to account for animal behavior could be established—insofar as the failure of any global scientific paradigm can be "established"—in a wide variety of experiments and in exquisite detail because of the theory's allegiance, in the end, to testing and evidence. In the case of psychoanalysis, it remains possible that improvement in measuring techniques, more accuracy in the fine-grained reconstruction of the past, a more complete reconstruction of recent events, and a more complete set of psychodynamic laws might greatly improve psychoanalytic prediction. We cannot gainsay this, except to say that no such improvements have yet emerged.

The Logic and Benefits of Expectation

The reluctant acceptance of expectations by behaviorists was no small concession. Properly understood, it opens the way to a fundamental reorientation in thinking about how past experience influences behavior—not through the direct molding of behavior but through information about possible futures. Choice now makes sense. Lashley's rats, even while confined to the narrow channels of the maze, appear to have been building up an evaluative map of the possibilities their environment afforded, stretching well beyond actual experience and enabling them to improvise opportunistically on the spot. Such behavior draws attention to another core aspect of cognition that is oriented toward prospection: the active, selective *seeking* of information ("exploration"), which, if we are right about prospection, should be as vital as the active, selective *processing* ("exploitation") of information (Rescorla, 1988).

So far we have argued that a long history of development in the empirical explanation of behavior points toward guidance by prospective representations—"if–then" possibilities. In this section, we offer an a priori argument for the centrality of expectation in current models of rational cognition and choice, and we then consider some striking evidence from ecology and neuropsychology that animals and humans might actually implement these models. This also enables us to contrast our approach to prospection with other views of prospection.

Consider how a systems theorist might approach the challenges facing a living organism. The good regulator theorem (Conant & Ashby, 1970; Eykhoff, 1994) suggests that for the brain to be a good regulator of interactions with the environment, both physical and social, it must build and use a model of that environment. Part of such a model will be if–then conditionals, both about what to make of incoming information and about what acts would have what effects. Energy-wasting effort and costly surprises as the organism makes its way in the world will be minimized if the organism is guided by an accurate model of pairings of the form: "*if* in circumstance *C* and state *S, then* behavior *B* has outcome *O* with probability *p.*"

The challenge, then, is without an expert engineer to design it, how could an organism acquire such a model? This would seem to be an impossibly complex task were it not that learning can take the following form:

$$\text{expectation} \rightarrow \text{observation} \rightarrow \text{discrepancy detection} \rightarrow \text{discrepancy-reducing change in expectation} \rightarrow \text{expectation} \ldots \quad (1)$$

Schema (1) describes a family of *feed-forward/feedback* models of learning and control familiar to engineers designing "smart" or "adaptive" systems—such as Web sites that "learn" your preferences by offering you options and using your choices to "improve" their offerings the next time. The feed-forward/feedback idea is at the heart of the prescient learning theory of G. Miller, Galanter, and Pribam (1960), of much adaptive control theory (Åström & Murray, 2008; see also Carver & Scheier, 1990), and of the Bayesian revolution in contemporary epistemology (Earman, 1992).

Expectation is pivotal in schema (1) because it transforms experience into experimentation—continuously generating a "test probe" so that the next experience always involves an implicit question and supplies an answer, which can then function as an error-reducing "learning signal." Empirically, error-based learning of this kind can be seen, for example, in the optimization of eye motion (Soetedjo, Koyima, & Fuchs, 2008) and muscular-skeletal motion (Scott, 2004; Todorov & Jordan, 2002). Such learning is a matter of responding not simply accurately but passively to environmental features, as when a skilled athlete's rapid eye movements predict the trajectory of a ball more rapidly and accurately than an amateur's (Land & McLeod, 2000; see also Abegg, Manoach, & Barton, 2011), but also of spontaneous, self-initiated eye movements optimal for the extraction of information relevant to resolving task-relevant uncertainty (Najemnik & Geisler, 2005). It appears that top athletes are not simply well trained but act through a continuously updated, largely unconscious or "implicit" action-guiding forward model of their situation and its possibilities (Yarrow, Brown, & Krakauer, 2009).

It is a feature of learning systems akin to schema (1) that they begin with a bias—an expectation for the future that does not assign equal initial likelihood to all possibilities. This would appear to be a defect, in comparison with "unbiased" or accumulative learning. Yet as Rudolf Carnap showed in his foundational work on confirmation theory, this seeming "bug" is actually a feature. "Unbiased" confirmation functions, which start with equiprobability for all events and then eliminate options only as actual experience excludes them, fail to learn differential expectations (Carnap, 1950). Schema (1)-like expectation-based learning systems have many advantages over pure associative learning (Gallistel & Gibbon, 2000) and possess four key a priori epistemic features: (a) they permit learning from experience in Carnap's sense; (b) other things equal, the influence of initial expectations will tend to diminish as experience grows, so that initial bias tends to wash out; (c) other things equal, the expectation value will tend with increasing experience to converge on the actual underlying relative frequencies in the environment, if there are such (Good, 1960; Jeffrey, 1953); and (d) where underlying relative frequencies are unstable, they can remain flexible in response and can learn from variance itself (Courville, Daw, & Touretzky, 2006).

From ideal to real

Schema (1)-like learning systems have the additional advantage that they are relatively simple to implement, thus making a good approximation of Bayesian learning available to animals with limited memories or incapable of higher-order self-representations (Behrens, Woolrich, Walton, & Rushworth, 2007; Sanger, 1989). "Animals run on batteries," the evolutionary ecologists tell us, and there is no recharging once they have run out. Energy spent exploring must be constantly offset by energy gained exploiting, so effectiveness in both, and efficiency in balancing them, is at a premium. Likewise is true of selectivity and efficiency in gathering information from the environment (Dayan, Sham, & Montague, 2000).

Naturalistic support for this comes from long-standing evidence of near-optimal foraging in species as diverse as birds and moose, showing sensitivity to evolving marginal gains, costs, and risks (for a summary, see Dugatkin, 2004). Foraging mammals have systems of neurons whose firing rates and sequences correlate with differences in the identity of stimuli, their intensity, the magnitude of specific positive versus negative hedonic rewards or food values, the relative value of a stimulus (e.g., deprivation versus satiation), the absolute value of a stimulus (e.g., physiological need), the probability or expectation of a given outcome, the occurrence of a better- or worse-than-expected predicted error, and the absolute risk and expected value of given actions (Craig, 2009; Grabenhorst & Rolls, 2011; Kringelbach & Berridge, 2009; Preuschoff, Bossaerts, & Quartz, 2006; Quartz, 2009; Rolls, Tromans, & Stringer, 2008; Schultz, 2002; Singer, Critchley, & Preuschoff, 2009; Tobler, Dougherty, Dolan, & Schultz, 2006).

A recent series of experiments with rats in T-mazes makes such evaluative prospection vivid. Research on spatial representation has found multiple interlinked neural systems for representing Tolman's (1948) hypothesized "cognitive map" and self-location (Ainge, Tamosiunaite, Worgotter, & Dudchencko, 2012; Derdikman & Moser, 2010; Langston et al., 2010). During REM sleep, rats trained in a maze repeatedly reactivate the neural "map" they have formed during training, with activation preferentially located in areas they visited less frequently during the day. Moreover, shortcuts begin to emerge in the neural map, passing through areas never visited (Gupta, van der Meer, Touretzky, & Redish, 2010; Ji & Wilson, 2007). The rat is using resources of internal simulation to redeploy elements of past experience efficiently and creatively, departing from associationism's predictions in a manner that would help explain how Lashley's (1929) escaped rats knew just which shortcut to take.

Prospective versus habitual control

Now consider a rat back in its T-maze. When it faces the choice point, activations in its neural map spread alternately down the two arms ahead of the rat's current location, mentally exploring and assessing what might lie in store down the two arms before making its turn. This pattern of forward activation reflects comparative reward experience and predicts choice (Johnson & Redish, 2007; Johnson, van der Meer, & Redish, 2007). Such active, ongoing prospection nicely illustrates teleological control—navigating into the future by considering future possibilities and electing action in light of the benefits and risks they promise.

If there are side channels in the maze that never contain food, the rat soon learns not to explore these mentally as it trundles rapidly past them toward choice points that matter (Johnson & Redish, 2007). This illustrates well the difference between prospective guidance and habitual or "overtrained" behavior. Components of action can pass from goal-directed to habitual control when they can be successfully repeated without need for evaluation of alternatives (Schneider & Chein, 2003). However, because habitual control is largely unresponsive to degradation of the value of outcomes (Killcross & Coutureau, 2003; Yin, Knowlton, & Balleine, 2004), successful goal pursuit involves an overall teleological organization of action: Spontaneous goal-monitoring requires an expectation-based forward model relative to which failures and conflicts can be detected, and when this happens control tends to shift back to the goal-based system (Isoda & Hikosaka, 2007). When we speak of teleological behavior, we thus do not deny that habits can play a part in subcomponents of action, but this role takes place within a flexible, value-based framework that is the key to the intelligence of the behavior (Diedrichsen, Shadmehr, & Ivry, 2009) and to the skill of expert behavior (Yarrow et al., 2009).

Such overall teleological organization of action makes evolutionary sense. We hypothesize that animals become more effective and more populous if they do not live simply in the present but rather continuously model what might lie ahead and proactively seek information (Dayan et al., 2000), allocate mental resources (Bissmarck, Nakahara, Doya, & Hikosaka, 2008), evaluate alternatives (Kennerley, Walton, Behrens, Buckley, & Rushworth, 2006), and select action (Rushworth, Walton, Kennerley, & Bannerman, 2004). The advantages of this sort of global organization, we suspect, has been a vital force in shaping animal perceptual, cognitive, affective, and motivational systems. It seems that today's foragers, the products of eons of selection for efficacy and efficiency, have evolved a highly efficient brain that learns surprisingly like a Bayesian and allocates scarce time and effort surprisingly like a rational investor (Quartz, 2009), foraging the world for value and information. Or perhaps this is not surprising after all?

Contrasting approaches and errors of prospection

The systems underwriting such functional rationality or optimality in learning and choice are heavily conserved in evolution and are present in yet more developed forms in humans. Neuroeconomists have recently emphasized this convergence between the formal demands of value learning and rational decision making, on the one hand, and neural architecture and processes, on the other, including the differences between habit-based and goal-based behavior (Rangel, Camerer, & Montague, 2008). Our emphasis on functional rationality in the philosopher's or economist's sense contrasts with earlier, seminal work on prospection. Prospect theory (Kahneman, 2011; Kahneman & Tversky, 1979) has proven enormously fruitful in the study of actual decision making, but it emphasizes departures from rationality, describing various "heuristics" and "biases" in reasoning and decision making. Likewise, the pioneering work of Gilbert and Wilson, who introduced the term "prospection" (Gilbert, 2006; Gilbert & Wilson, 2006, 2007), also emphasizes failures and errors in "affective forecasting." Moreover, Gilbert and Wilson see simulation as a top-down cortical process, necessarily episodic because conscious mental activity has other things to do. They think of prospection as providing only "cardboard cut-outs of reality" that briefly "trick" subcortical structures into responding as if to actual sensory experience (Gilbert & Wilson, 2007, p. 1353). According to our hypothesis, generating simulations of the future can be conscious, but it is typically an implicit process—not requiring conscious initiation or monitoring, often not accessible to introspection, and apparently occurring spontaneously and continuously. Indeed, even when individuals engage in conscious prospection, their intuitive sense of the value of alternatives may be underwritten by unconscious simulation (Railton, in press).

Unlike conscious processes, the unconscious processes underlying implicit prospection are capable of handling very large numbers of statistical relationships at once—think, for example, of the decisive feint and pass made at the last minute by a champion soccer player, setting up the winning goal. At the conscious level, Kahneman and Tversky (1979) showed, humans are poor at explicitly calculating conditional probabilities. But even 8-month-old children appear capable of rapidly learning conditional probabilities implicitly, as they spontaneously parse the speech they hear (Aslin, Saffran, & Newport, 1998).

Recent years have seen considerable attention to unconscious processes, which are often seen as "automatic" and coarse-grained in their responses (Bargh & Chartrand, 1999; Bargh & Ferguson, 2000, Haidt, 2001). By contrast, the implicit processes of prospection we have emphasized appear to be selective, flexible, creative, and fine-grained. They fit into a view in which subcortical affective processes are integral with cognition (Pessoa, 2008) and serve to "attune" the individual intelligently and dynamically in response to a changing environment (Ortny, Clore, & Collins, 1988; Reynolds & Berridge, 2008; N. Schwartz, 2002; Schwartz & Clore, 2003).

The logic of expectation as studied by control theory, confirmation theory, and rational choice theory, we have argued, provides normative support for a model of the mind organized centrally around the tasks of forming and revising if–then act–outcome pairs, separately encoding risk, reward magnitude, and expected value to provide a basis for choice in the face of uncertainty. We contrast our view to the emphasis on the errors of prospection. Mindful of the fact that not getting it wrong does not remotely equal getting it right, we emphasize how accurate prospections can be. Much prospection appears to share the architecture of the optimal models developed a priori in philosophy, economics, and systems analysis. This picture helps us understand how animals are able, through schema (1)-like learning systems, to exhibit impressive approximations of optimality. And in conscious beings such as humans, implicit processes can help explain the success of more conscious decision processes that are attentive to "intuitive" evaluations. Such processes help keep human decisions rooted in reality and are largely effective—as they must have been to explain humanity's evolutionary success.

To be sure, implicit statistical learning processes are vulnerable to small samples, unrepresentative samples, and the limitations of temporal feedback. These can lead to distortions in prospection and failures to anticipate accurately. Human consciousness, discussed below, permits statistical learning to be supplemented with supervised learning and direct social learning, a tremendous advantage. So we do not wish to minimize sources of error or to fail to acknowledge the role of explicit reasoning processes in helping to counteract such effects. But we believe the time has come for a shift in emphasis when discussing prospection away from error—informative and fascinating as this has been—and toward success. Moreover, such a shift toward the positive could have a salutary effect on understanding the errors themselves. We note with interest an emerging literature in which various familiar heuristics and biases are accounted for by adaptations that could figure in optimal learning strategies (Denrell, 2007; Le Mens & Denrell, 2011; March, 1996).

Desire versus drive

We have argued that effective learning is expectation based, and so this should apply to motivation as well as cognition. This is the critical difference between *drive* and *desire.* Theorists as diverse as Freud (1920/1975), Lorenz (1966), and many behaviorists (e.g., Dollard & Miller, 1950; Hull, 1943a) placed the concepts of *drive* and *drive reduction* at the core of motivation. A drive was a motive force, arising from deprivation and impelling behavior that reduces this discomfort. Moreover, this reduction reinforced the behavior—so no goal or telos is needed. This wonderfully simple idea, Berridge writes, "is so intuitive that it was thought to be self-evident for decades . . . [and] some behavioral neuroscientists today still talk and write as though they believe it. All the more pity, perhaps, that the idea turns out not to be true" (Berridge, 2004, p. 191).

Behavioral studies first cast it in doubt. Animals whose nutritional needs were met intragastrically retained a lively interest in eating (N. E. Miller & Kessen, 1952; Turner, Solomon, Stellar, & Wampler, 1975). Later, brain stimulation studies showed why—electrical brain stimulation producing eating is not aversive, as a drive concept would have it; it is a reward (Berridge, 2004). As everyone knows intuitively, eating is attractive to contemplate—an object of *desire*—quite unlike forcing one's hand into ice water to escape the pain of a burn.

Philosophers since Aristotle have emphasized that desire is not a blind urge but rather represents its object as an "apparent good" (Aristotle, ca. 330 BC/1999, p. 1113a15) or under a "desirability characterization" (Anscombe, 1957, p. viii), an attractive prospect that can elicit motivation to seek it—"liking" a representation gives rise to "wanting" its object

(Railton, 2002; on the distinction between "liking" and "wanting," see Berridge, 2004). Novels, poems, and plays are written around desire (not drive reduction), because desire provides an intelligible teleology for human action, a narrative arc stretching from discovery of a transfixing but distant prospect, across the drama of longing, seeking, and overcoming obstacles, to arrive at a denouement in union with the object of desire.

Are we saying drivelike motivation never occurs? No. Addiction and salt deprivation, for example, can produce wanting without liking (Robinson & Berridge, 2000; Tindell, Smith, Berridge, & Aldridge, 2010). Certain physiological demands, natural or artificial, can produce "driven" motivation even in the face of profound distaste and resistance, but this is atypical indeed. Ordinary action, even eating a meal when hungry, does not work this way—for hunger makes eating attractive, not distastefully compulsive.

The difference between desire and drive speaks to another subtle but theoretically powerful issue, namely, active versus passive processes in choice. The driven-by-the-past framework makes agency and choice difficult to understand—individuals are responders rather than navigators. The past, of course, cannot be changed, so the very forces determining one's behavior are always out of one's hands. If instead we see the individual as using past experience as information, as continually forming and evaluating a range of future possibilities, and as electing action from among these possibilities in light of what she likes and values, then we can see that active agency is a natural part of the causal structure of action. Motivation for such action is not determined by fixed drives or past conditioning but is elicited by the evaluative process itself through the normal working of desire.

Not to be forgotten, however, are the less *agentic* aspects of simulating and evaluating possible futures. Some futures are consciously considered and evaluated, while others, perhaps most, occur spontaneously. Our title emphasizes *navigating* the future because, like a navigator, the organism must not only act but also mentally explore options and keep track of progress. This can be done by deliberate calculation and monitoring, but such effortful activity tends to be slow and to divide attention. Prospective guidance thus also includes spontaneous cognitive and emotional activity: intuition, undirected recollection, mind wandering, mental intrusions, creative inspiration, uneasiness, surprise, and satisfaction.

We will return to the topic of human agency in the final section of this article.

The Prospecting Brain

We have now presented two prongs of the argument that prospection is required for psychology: one historical and empirical—that animal learning could not be successfully explained without teleology; the second logical and evolutionary—that effective and efficient adaptive learning and regulation of behavior favors a teleological organization of behavior. We now present a third prong—the prospecting brain.

Prospection is at the core of four kinds of mental simulations (Buckner & Carroll, 2007).

The first is literally navigational: Imagine your home and the nearest supermarket. Now mentally walk block by block from your front door to the nearest supermarket. If you turned right out of your front door, now do the exercise again starting with a left turn and taking therefore a different route to the supermarket. (If you turned left, start by turning right.)

The second is social, about other minds: Imagine that you have been invited to have a chat with President Obama about whom he should choose as the "White House Person of the Year." Whom would you nominate? How would the president react to that nomination? Now imagine telling the president three reasons why he should choose your nominee. Imagine the objection that you think he would most likely raise. Now imagine your response to that objection and how effective it would be for the president.

The third is intellectual: As you are reading this article, with its grand claim of the central organizational role of navigating the future, what mental activity are you engaged in? If you are like most active readers, you are mentally trying out various reactions to the material. You are making arguments against the idea, or finding holes in our reasoning or weaknesses in our evidence, or thinking how you might improve, qualify, or defend the article's conclusions. You are imagining trying to explain the argument to someone else, perhaps a class or a colleague, or how you might use it to advance your own positions. You are asking whether the idea is really so radical after all, or you are building toward a decision not to waste any more time on this article.

The fourth exercise is memorial. Recall a happening in your life that turned out badly because of something you said or did. What could you have said or done that might have made it turn out better? Run through that scenario. How would things have been better? Really? What would have been the negative consequences as well?

What the first three exercises demonstrate is the process of prospection, the mental running of hypothetical simulations of the future. What the fourth demonstrates is the mental running of counterfactual simulations of the past. Each of these exercises demonstrates your enormous facility for generating, exploring, and evaluating alternatives to the present you now confront or the past you cannot help but remember (Buckner & Carroll, 2007). In each case, you free yourself from your actual conditions to take advantage of your powerful mind to "do," explore, assess, and perhaps learn from mere possibilities.

What evidence is there that these four kinds of mental simulations share something in common? Intriguing new evidence suggests there is a single core brain network involved in these diverse forms of imaginal simulation. Evidence for this hypothesis is closely linked to the discovery of the *default network*, a set of interconnected regions implicated in internally directed mentation. We now discuss key findings that led to the discovery of the default network and established its central role in prospecting future events.

Neuroimaging studies are typically structured with alternating time blocks—blocks of task-related activity alternating with resting blocks, which serve as a control condition. Contrasting levels of brain activation during a task block against a resting block reveals neural activation specifically associated with the task. Researchers had long known that the reverse contrast in which the resting block is compared with the task block shows a highly reliable pattern of neural activation in the brain's midline and the lateral parietal lobes (Shulman et al., 1997), but the significance of this finding was not immediately evident. One might have thought that when subjects are given no explicit instructions (other than to look at a blank screen), the resulting patterns of brain activation would be highly unstructured and thus would be very noisy. Yet these studies show that the resting state is associated with a highly reliable and uniform pattern of activation. What might be the significance of this pattern?

Marcus Raichle, Debra Gusnard, and colleagues, building on suggestions of other theorists (e.g., Andreasen et al., 1995), proposed that the regions exhibiting enhanced neural activations at rest relative to task constitute a functionally integrated neurobiological system (Gusnard, Akbudak, Shulman, & Raichle, 2001; Raichle et al., 2001). Using positron-emission tomography (PET) imaging, they showed that the pattern of enhanced activations observed during the resting state are best interpreted not as transient deviations from a quiescent baseline but rather as a return to an active state involving ongoing tonic brain activation (Raichle et al., 2001). They called the regions tonically activated at rest the "default mode" system and proposed that these regions support internally directed mentation during intervals when there are no externally cued cognitive demands.

Parallel investigations assessing *functional connectivity* (i.e., patterns of correlated activity between brain regions) provided important convergent support for the default mode hypothesis. It has long been known that neural activity exhibits a pattern of spontaneous, slow oscillation (<0.1 Hz) in individual neurons and larger collections of neurons (Biswal, Yetkin, Haughton, & Hyde, 1995). More recently, researchers discovered that various methods (primarily functional MRI imaging but also PET and electro- and magneto-encephalography) can be used to quantify temporal correlations between slowly oscillating neural activity across brain regions. Functional connectivity methods demonstrate that the collection of regions more active at rest (i.e., default mode regions) constitute a highly coherent large-scale network. These regions exhibit patterns of functional connectivity that are stable across conditions including during rest (Greicius, Krasnow, Reiss, & Menon, 2003), during active tasks (Laird et al., 2011; S. M. Smith et al., 2009), and over time (Shehzad et al., 2009; Zuo et al., 2010) and which correspond to biological markers of network structure such as anatomical white matter tracts (van den Heuvel, Mandl, Kahn, & Hulshoff Pol, 2009).

The default mode regions are thus properly thought of as a highly interconnected network, but what activity are they organized around?

Dozens of neuroimaging studies have assessed the neural basis of tasks involving the four forms of imaginal simulation above. Quantitative meta-analyses of these studies (e.g., Spreng, Mar, & Kim, 2009) have demonstrated that there is substantial overlap in the regions implicated in these tasks, suggesting that a common underlying network subserves these functions. Moreover, the regions implicated in these tasks correspond to the regions identified by Raichle, Gusnard, and colleagues as part of the default mode system. This supports the intriguing idea that when people are at rest and not engaged in some externally directed task, their mental lives are largely occupied by imaginative simulations (Buckner & Carroll, 2007; Mason et al., 2007). This hypothesis gains additional support from two recent studies. The first found that people with a more pronounced tendency to engage in internally directed mentation (such as daydreaming) generated more robust activity in the default network during rest intervals of functional MRI (fMRI) scans (Mason et al., 2007). The second study used ecological momentary sampling during fMRI scanning of a rote cognitive task and found enhanced activity in the default mode network during intervals when subjects' minds wandered off task (Christoff et al., 2009).

Remembering the past and imagining the future

The default mode hypothesis set the stage for research on how these various forms of imaginative simulation are interconnected (Buckner & Carroll, 2007; Goldman, 2008). One particularly fruitful line of investigation has characterized the rich connections between episodic memory of the past and prospective representations of the future.

Episodic memory refers to autobiographical memory for specific prior events, including information about who was present, what occurred, and what was felt (Tulving, 2002). Episodic memory appears to be a fundamentally *constructive* process. Each time an event is remembered, past episodes are reconstructed anew and in most cases a little bit differently than the last time. At first it was thought that this reflected an initial process of memory consolidation, by which memories had to be replayed several times until they stabilized, after which each remembering would be more or less identical. However, more recent work has concluded that memories continue to change and evolve as long as the person lives. Memory never "nails things down" once and for all. This has puzzled researchers for decades (Bartlett, 1932; Neisser, 1966). Why is memory designed in such a fallible way? According to the constructive episodic simulation hypothesis put forward by Daniel Schacter, Donna Addis, and their colleagues (Schacter & Addis, 2007a, 2007b; Schacter, Addis, & Buckner, 2008; see also Szpunar, 2010, for a detailed review), episodic memory provides details needed to construct prospective simulations of future events, and thus the two rely on some common substrates. Both episodic memory and prospective simulation rely on such common processes as the storage and recall of individual details, mental imagery, and self-referential processing. In addition, both involve constructive operations that bring together these elements into a coherent episode.

Prospection, however, engages these constructive processes more vigorously, as the task of constructing multiple possible futures requires extrapolation beyond the given evidence to what would be the case under hypothetical and even counterfactual conditions. Support for Schacter and Addis's hypothesis comes from recent neuroimaging studies that identify a core network (which encompasses the regions of the default network discussed earlier) that is engaged during both retrospection and prospection (Schacter & Addis, 2007a; Spreng et al., 2009). Prospection preferentially engages certain regions that are required for flexible recombination of information, including anterior hippocampus and frontopolar cortex (Addis, Cheng, & Schacter, 2010; Addis, Pan, Vu, Laiser, & Schacter, 2009; see also Szpunar, Chan, & McDermott, 2009).

Other cognitive capacities also engage the core network identified by Raichle and colleagues. One is reasoning about counterfactuals (Van Hoeck et al., 2012), which are if–then statements in which the antecedent is known to be false (e.g., If I had not taken that job, then we would never have broken up). A second is reasoning about other minds (Gallagher & Frith, 2003).

Philosophers and psychologists have long posited that simulation plays a central role in these two psychological abilities (Goldman, 2008). Reasoning about counterfactuals is proposed to involve one's imagining a world as similar as possible to the actual world but for the fact that the antecedent of the counterfactual is true. One then assesses whether in this simulated world, the consequent is true (Byrne, 2002; Williamson, 2008). And simulation provides a route to understanding what is in the hearts and minds of those around us. Brain-imaging studies indicate that simply observing the behavior, discomfort, or disgust of another person can activate corresponding regions in one's own brain (Decety & Chaminade, 2003; Lamm, Batson, & Decety, 2007; Wicker, Keysers, Plailly, Royet, & Gallese, 2003). Such apparent simulation also arises spontaneously when contemplating one's own future, and individuals with profound deficits in this capacity have difficulty with both prudential and social behavior (Demurie, De Corel, & Roeyers, 2011). Humans are extraordinary among animals in their capacity for the prospection necessary for long-term, shared, stable enterprises such as government, law, schools, commerce, collective bargaining, and retirement planning. Commitments, relationships, values, and convictions, even "the self" as a persisting entity, are a matter not just of how one has acted or is acting but of how one thinks about the future and how one will or would act in various futures (see Northoff et al., 2006). Mental simulation enables one to envisage these futures and to think through one's responses to them before they arise. In this way, ongoing prospection is an important part of whom one is and how one stands among others.

Buckner and Carroll (2007) conjecture that the fundamental function of the core neural network implicated in episodic memory, thinking about the future, counterfactual reasoning, mind reading, and spatial navigation is simulation. An ability to project oneself into other times and other shoes could have held an evolutionary advantage strong enough to shape neural architecture itself, and the existence of a core brain network that supports prospection and other forms of imaginative simulation constitutes the fourth and final prong of our argument.

Prospecting Prospection

We intend the idea that intelligent action is guided by assessment of future possibilities rather than driven by the past to be a framework permitting the integration of a wide range of research, from animal learning theory to rational choice theory, from empathy and emotion to motivation and control, from remembering the past to projecting the future. We believe that the idea that a prospective orientation is fundamental to mind will be more fruitful at predicting and understanding action than was the framework in which action is explained as solely driven by the past.

One way of judging the fruitfulness of a new framework is whether it illuminates major issues that were opaque in the framework it seeks to replace. We suggest that taking prospection seriously might reorient thinking and illuminate (a) consciousness and subjectivity, (b) free will, (c) and psychopathology and therapy.

Why consciousness?

We speculate that a major function of human consciousness is to permit better prospection of the future. Consciousness enables better individual prospection and better shared prospection. Tomasello (1999) noted that although chimps may hunt as a group, each individual essentially acts on his own, looking for individual opportunities. Such behavior exhibits a failure to solve the "Stag Hunt" problem, first identified as essential to understanding human sociability by Jean-Jacques Rousseau (see Skyrms, 1996, 2004). To solve this problem, it must be possible to coordinate in a distinctive way, extending over time. The group must share information, plan together, and jointly carry out these plans, monitoring each other's progress and adjusting accordingly. Each of these tasks moves individuals out of full preoccupation with their own current situation, and each calls for extended, shared prospection. We speculate that consciousness is indispensable for the extraordinary scope and effectiveness of this ability in humans—an ability that might, as much as any other, explain how physically feeble humans became such a potent force in the world.

Baars (2002) reviewed evidence that consciousness of the sort that we call prospection is crucial for *sequences* of thought. Priming studies that rely on unconscious processes have generally failed to work even for two-word combinations (e.g., Draine, 1997; Greenwald & Liu, 1985). In dichotic listening tasks, people can process and respond to single words in the unattended channel, but meanings of sentences do not get through (Mackay, 1973). Logical reasoning—of which coherent planning for hypothetical situations is an example—requires directed construction of sequences of thoughts and directed imagining (De Neys, 2006; DeWall, Baumeister, & Masicampo, 2008; Gilbert & Wilson, 2007; Lieberman, Gaunt, Gilbert, & Trope, 2002). A shared space is needed for such construction, where sequential assembly can take place under the guidance of means–end and subplan consistency—prospective processes requiring management of attention and stepwise, concept-based inference. Baumeister and Masicampo (2010) proposed that conscious thought can be understood as such a place (see also Smallwood, Nind, & O'Connor, 2009; Smallwood et al., in press).

Social discourse is also such a place. Although many complex acts can be executed without conscious supervision, talking coherently is not one of them: People have to be conscious of what they are saying. The mental processes that produce conscious thought are closely linked to those that produce speech, and in fact M. C. Fox, Ericsson, and Best (2011) showed that having people voice their thoughts as they perform a task usually has little or no effect on performance, which suggests that talking does not constitute much extra mental load over and above thinking one's way through the task. To reason and plan together—working out what to expect, or feel, or seek under a variety of contingencies and comparing one's own responses with others'—individuals need to be able to survey and convey to others their own thoughts, feelings, memories, and imaginings. This requires not only representational capacities, for which consciousness is surely not needed, but meta-representations that we can make publicly available—including embedded if–then meta-representations of what one would think, feel, or seek to do under hypothetical circumstances. Were our own thoughts, feelings, memories, and imaginings a blank to us, or were we unable to acquire a model of ourselves through the experience of them, we would be confined to the sort of signaling found in animals. Such signaling appears to be useful for manifesting current emotional states or attention and thus conveys valuable information about what to expect next. But it would also appear to be profoundly limited in its capacity to share direct information about mental states such as episodic memories, conditional commitments, beliefs of a general nature, or beliefs about what might be the case, or would likely be thought or felt, at some distant future time. An extensive literature review by Roberts (2002) concluded that most animals can project at best 20 min into the future. The adaptive advantages of shared planning, then, would have favored mental economies with greater *access consciousness* (Block, 2008), the wide availability of information and attitudes for a variety of mental operations, including inference, deliberation, planning, and voluntary action.

We do not know how the snowball of language and culture first began to roll. But perhaps it grew so dramatically and quickly because consciousness and communication enabled humans to bring the processes and products of prospection into the light. Language and culture are multipliers of the effectiveness of prospection given that many minds are so often better than one—creating a wider pool of evidence, shared imagination and examination of alternatives, functional specialization, and coordinated responses. So our speculation is that consciousness makes for better prospection—and better sharing of prospection.

Why subjectivity?

This suggests a different way of thinking about the question, "Why do we have subjectivity at all?" What does subjectivity—the phenomenology or the "felt qualitative character" of experience itself, *phenomenal consciousness* (Block, 2008)—add to whatever can be accomplished without it? Why does conscious prospecting also have a subjective face? Our only excuse for speculating about this aspect of the "hard problem of consciousness" (Chalmers, 1996) is that there are not many other plausible answers.

Subjectivity might provide an effective competitor to one's own current experience—so past episodes and future possibilities can compete with and be compared with ongoing experience and can have comparable effects on feeling and motivation. How can a particular lesson selected from the past have full impact in thinking about the future or a distant goal be imbued with the salience and attraction needed to sustain goal pursuit over the long term? Vivid imagination could provide the salience and appeal needed if, say, a personal ideal of becoming thin is to compete successfully with a present rich dessert or if attacking an enemy is to shine as an ideal to compete successfully with the more proximal fear of harm.

Prospection will be an effective competitor to actual experience only if it can be equally affecting. Consider the problem of staging an affecting play. Here one must provide a convincing artificial experience, and this is done by creating powerful appearances to reproduce the impact of the real. Appearance—a subjective state—is what ongoing, imagined, and remembered experience have in common, their common coin. Human imaginative prospection is like the staging of a play using appearances to simulate the actors and events; but it must be more, for one must be in the drama, and the script must feel real.

Consciousness and commensuration

Consider a concrete example of self-conscious prospection. (We ask the scholarly reader to forgive our venture into phenomenology here, but the point is difficult to make without an example that illustrates how very many evaluative simulations of possible futures occur in a brief time.) Martin Seligman is sitting at his computer in the process of deciding what to do in the next hour. Here is what he was conscious of in the course of about 30 seconds:

1. I could go onto the Internet and play bridge. Who might be available? Mark Lair, but he's usually at lunch. Peter Friedland, he's in Taiwan, probably going to sleep, but there are a few lesser lights, likely to be available. They make errors. Anyway, I've wasted a lot of time playing bridge lately.
2. I might help Mandy teach Carly and Jenny about the Silk Road. I don't know much about Asia. But the kids would love it, and I haven't spent any time with them today. Mandy might find it intrusive, having prepared the lesson. But she thinks I have not done my share of teaching lately.
3. I might make myself some lunch. There's some Moroccan chicken left over in the fridge. It's pretty high calorie though. And I'm meeting Phil Voss for dinner at Le Bec Fin in only 5 hr. But I could order only their three course meal. Maybe Mandy was saving the chicken for the kids' dinner.
4. I could keep working on this damn article. But I'm having trouble thinking through examples of compelling counterfactual simulations. Maybe a bridge break will help. But this is itself a pretty good example, so maybe I should keep plugging. Why bother? I don't have a deadline, since this article is for my own amusement. Peter will be disappointed if I don't follow up soon.
5. All those tulip bulbs need planting. I could use the exercise, particularly with Le Bec Fin coming up. The temperature is good, but the ground is soggy. Tulips can be planted even if it gets really cold, no rush. They might rot. I did lift weights for 20 min already today. But I could use some fresh air. It would calm me down. I need it particularly after the argument with my dean.

Notice how multidimensional these simulations of the future are, notice that each is evaluative, and notice also how incommensurable the evaluations seem. By what metric does the pleasure of playing bridge with Mark Lair stack up against annoyance of letting tulip bulbs rot or the satisfaction of seeing brightly colored tulips in 6 months or the anticipation of Moroccan chicken or the guilt of not working out? In the market, disparate human actions achieve commensurability through the unifying metric of money: so fixing a hole in the roof and a loaf of bread become commensurable by attaching a dollar value to each and computing how many loaves of bread you can buy in exchange for fixing your neighbor's roof or how many loaves of bread you would have to forgo to have your own roof fixed.

Affect is the brain's common currency for value, and conscious, subjective affect would permit the possible futures to be brought into the open for explicit comparison with each other. We have argued that conscious subjective affect attached to prospections would enable them to compete effectively with ongoing experience. Here we are asking how such prospections can compete with one another. Each one of Seligman's competing prospections has a (perhaps nebulous) global affective valence, but within each there are also conflicting valences (he does not know much about Asia, but the kids would love it, and Mandy thinks he has not done his share). Properly responding to many situations, and especially social situations, is not a matter of feeling one particular shade of one particular emotion. On the contrary, taking the full measure of the situation will involve keeping before the mind a mixture of "unblended" feelings of different shades and degrees, sometimes even conflicting feelings representing diverse perspectives that should not be combined or resolved. Nonconscious summing methods may be adequate when the values at stake and appropriate responses are straightforward, even if complex—as in optimal foraging or operating from a fixed set of preferences over multiple dimensions. But what if certain foods are taboo and yet one's need is desperate and one's elders disapprove yet one's children are at risk and more powerful neighbors look down on anyone who respects this taboo? One will need to act, but doing so successfully is not something easily passed along to an algorithmic learning system that produces a net action tendency. And in such cases there is as much a question about how to feel or what to show as feeling as how to act. Often every available act has costs, so that even choosing the best act may also require making amends to those disadvantaged by it or steeling oneself for their disapproval.

Given multiple incommensurable dimensions and conflicting values and perspectives, none of which will go away even if a sum could be struck, we do not quite see how implicit algorithms could do the job in real time (see also Morsella, 2005). If they could instead feed into a final common but nonsummative path, that is, an experientially rich and detailed workspace, then perhaps humans would be able to use their intelligence and imagination to best effect. Human life might be much the same if humans saw in black and white rather than color or saw the colors differently. But it would be incalculably different if humans could not keep before their minds the often conflicting thoughts and feelings and memories afforded by experience or use voluntary thought processes to consciously prospect alternatives in light of these "unblended" or "unsummed" facts and values. To be sure, in daily life and in most of the regulatory decisions made by the executive brain, there are genuine constraints of cognitive resources and of time. Consciously keeping track of all the component elements would be out of the question. But even in such cases, it can be best to act in awareness of complex feelings and conflicting thoughts—indeed, this seems to be the kind of thing that humans specialize in.

Freedom of the will

Philosophers have traditionally posed the question of free will in highly abstract metaphysical terms: If a universe obeys deterministic laws, can agents in this universe be free? Compatibilists argue that if one is careful not to confuse *causation* with *constraint*, then there is no opposition between free will and determinism, while incompatibilists deny this claim. By all accounts, this metaphysical debate has reached a stalemate. There are a variety of compelling arguments on either side, but nothing decisively tips the balance in one or the other way.

We are not going to enter the metaphysical fray about determinism in this article. Rather, we see a need to develop additional perspectives on free will that might move us past the deadlock. One approach that has become increasingly popular within philosophical and psychological circles locates a more modest, distinctively antimetaphysical point of entry into the free will debate (Baumeister, Crescioni, & Alquist, 2011; Mele, 2001). The basic idea behind this approach is, very roughly, to put questions of psychology ahead of questions of metaphysics, and here is how it works. Start with the question of what component psychological mechanisms or capacities a creature needs to have in order to be free and autonomous. Build a catalog that encompasses the full assortment of "design features" that make an agent free, using psychological and neuroscientific data from humans and other animals as critical guides to theory construction. Later, after the catalog is complete and the functional specifications of the items on the "free will inventory" are fully fleshed out, then, and only then, are abstract metaphysical questions broached. In particular, questions can then be posed in a more structured and informed fashion about what kinds of physical laws, properties, or substances must exist in the universe (e.g., irreducible chanciness, certain forms of agent causation) in order for the capacities on the list to be realized. We call this the *capacities-first* approach to free will, and we believe that prospection science provides a theoretical framework for organizing this kind of inquiry.

What kinds of psychological processes appear to be implicated, when we take ourselves to be acting freely? Stillman, Baumeister, and Mele (2011) sought to discover how ordinary people understand free will. They asked participants to narrate an event from their lives in which they acted of their own free will or, in another condition, not of their own free will. Actions reflecting free will were more likely than the non-free-will actions to emphasize pursuit of long-term future goals. Free actions were more likely than the nonfree ones to be about conscious deliberation and reflection. The free actions were also more likely to be consistent with the person's moral values, and the free actions were also more likely than unfree ones to bring about positive outcomes. Taken together, these findings show that everyday understandings of free will are about long-term, beneficial outcomes, aided by conscious reflection and principled commitments.

Freedom. These features point to the centrality to the experience of freedom of guidance via prospection. Acting freely involves the absence of constraint, but it is also fundamentally a matter of generating and evaluating multiple possible future courses of action and electing an act in light of them. It follows, perhaps, that enhancing freedom will involve enhancing the power of generating and evaluating options. Three distinctive design features of human prospection expand the complexity, time horizon, and accuracy of prospection.

Complexity. Plans consist of sequences of actions linked in a coordinated way to achieve a goal. Plans are usually decomposable into parts, each of which achieves some proximal subgoal (G. Miller et al., 1960). In order to get to Boston, I need to fill the tire with air. Once that is done, I can drive to the gas station. With the tires filled and the gas tank full, I can drive all the way to Boston. Often parts of plans are themselves decomposable into further parts.

Given their decomposable structure, efficient construction of plans requires a distinctive kind of prospective ability. It is not enough to simply prospect the outcomes of single actions. Rather, one must be able to perform *sequentially linked prospections*, and the attendant ability to rearrange and build more flexible plans not only expands the size of option sets but makes available a vast array of new options that are dramatically more likely to achieve one's goals.

Time horizon. Some actions unfold over a short duration: The rat in a T-maze selects going either right or left and receives a reward in a matter of moments. Humans, unlike other animals, can project years ahead and adjust current behavior accordingly. This contrast between humans and rats is set out not to deride the achievements of animal minds. When at a fork in a maze, expecting that danger lurks down one path and not another is an impressive feat. How much more impressive is it then that when at a fork in life, let's say choosing a major in college, we not only have the ability to prospect the immediate consequences of each option, but we can mentally "see" in rich and vivid detail the various ways one's life might unfold.

Accuracy about the future. Improvements in accuracy about the future include being able to imagine oneself in different subjective states that one will experience in the future. Human prospection includes representations of states and motivations that differ from one's current condition. If the anticipated motivation of his future self is one that the individual does not want to have, he can instead formulate options that prevent these anticipated future desires from arising. In one of the most famous works on free will in the modern philosophical literature, Harry Frankfurt (1988) argued that freedom of the will consists in having the will that one wants to have. According to Frankfurt, both humans and animals have the ability to have "first-order desires," that is, desires directed at doing this or that action. But humans also have the ability to step back and form second-order desires, desires about which first-order desires one wishes to have.

So specific design features of prospection expand freedom in at least three ways. Sequential prospection enables complex, flexible planning, expanding the number and quality of options. Prospection with a long time horizon enables options that unfold not just over days but rather over years and decades. And prospection with meta-representation enables better accuracy about the future.

Willing. We now turn from "freedom" to "willing" in order to explore the mechanisms by which a particular option is selected. There is one sense of "will" that comes into play when we engage in the spontaneous or deliberate prospection of future possibilities. This feels "free" because the mind freely explores possibilities, and it feels like "freely willing" because what precipitates our action is the making up of our mind among these alternatives. Nothing more, no additional act of will, is required to act "as I see fit."

So the experience of "freely willing" is running through these prospections until one feels that one's mind is made up and then taking the course of action one has settled on, and nothing more.

The "settled outcome" is in an obvious sense one's own idea, because it came about through one's own unimpeded mental activity, without internal compulsion (which is insensitive to what one prefers) or external coercion (which prevents one from weighing options without interference) or overpowering temptation (in which case the agent does not have the will he wants). No transcendental will is needed for the act to be "of one's own accord"; no rational homunculus must "freely endorse" it—for when the agent settles his mind after freely exploring options by following what "seems best," then if the agent also wants to have this be the ground of his choice, that is the agent freely endorsing it in every relevant sense and performing the act because of his endorsement. Indeed, were we to describe the activity of such a homunculus, it would be no different from this.

This description fits the phenomenology of freely willing. Had the agent thought more highly of another possibility, or had he been attracted to it whimsically and wanted his whimsy to guide him, he would have settled on that one. This is what "I could have done otherwise if I'd wanted to" means. This, moreover, is a notion of free agency worth wanting—because it enables us to pursue what we want. And one need only reflect on what one's life would be like were one beset by compulsion, serious addiction, or coercion to see what a difference free agency in our sense makes.

Our formulation is far from a complete account of the process of willing, and we are especially aware of two gaps in our account. The first is "how is one's mind made up" in leading to a "settled" evaluation among the options. In the section above on consciousness and commensuration, we postulated that a complex evaluative process among alternative options must exist, but we made few claims about the details of this process. In particular, we have assumed nothing about the role indeterminism or stochasticity might play (Glimcher, 2011; Grabenhorst & Rolls, 2011; Neuringer, 2002). But once prospection is put at the core of free will, this now becomes a tractable, empirical issue amenable to the traditional methods used in the fields, such as judgment and decision making.

The second gap is the role that "controlled" versus "automatic" processes play in prospecting. Under some conditions, external circumstances frame the question that is prospected. "If you had ten million dollars to spend over the next year, what would you do?" sets off a panoply of automatic prospections without need for any voluntary action ("a clinic in North Philadelphia, no, in South Sudan."). Under some circumstances, an effortful, controlled process frames the question that is prospected: "I have two deadlines this week; how should I meet them?" Under some circumstances, an internal state such as thirst cues the process. It is clear that a mixture of automatic and controlled processes are involved in the initiation, maintenance, and conclusion of the prospecting process, and the best we can say is that this now becomes a tractable and empirical issue.

The central point of our analysis of willing is, however, that there need be posited no such thing as a "will." It is worth remembering that the modern notion of a "will" as a thing is in fact a modern reification. Aristotle's term *boulesis*, used for the desire that combines with an idea of an act to yield rational action, comes from the notion of "taking counsel" or "thinking over," a form of prospection, and for Aristotle, decision (*prohairesis*) is "deliberative desire," that is, a desire to do something here and now activated through deliberative assessment of available acts, not an inner act of willing (ca. 330 BC/1999, pp. 1112b26, 1139a21–b5). In ancient Greek thought and law, the equivalent of "willing," ἑκών, meant being favorably disposed to seeing one's idea of an act brought into being; and murder, φόνος, was a matter of foreseeing the death of another, wishing for it, and doing it as a result. In Old English and in many contemporary Germanic languages, "I will it" is synonymous with "I like it" or "I want it"; and in many Latinate languages, "willfulness" (e.g., *volonté* in French) has the same root as "desire" (*vouloir*). "Free will" for the French is *libre arbitre*, roughly, "free weighing and judging," with no reference to a special volitional faculty.

Ultimately, our capabilities approach to the free will problem will face the question of what *metaphysical* assumptions, if any, are needed to fulfill the "job description" of free agency. It could turn out that our universe does not meet these requirements. But as far as we can see, the phenomenology of free agency can be understood, the psychological processes that underlie it described, the exceptions to it diagnosed, and the considerations that make it valuable to us illuminated without challenging anything in current physics.

Therapy and psychopathology

The question of whether we act in light of future prospects rather than being driven by past forces is not just academic. The major psychotherapies invented in the 20th century were founded on the premise that humans are driven by the past, and so the therapies naturally focused on undoing the dirty work of the past. Psychoanalytic theory holds that present symptoms are caused by unresolved sexual and aggressive conflicts from the past, usually the distant past. So therapy focuses on reliving past events and gaining insight into them. It is not an overstatement to say that the results of 100 years of this kind of therapy are disappointing.

Behavior therapies, similarly, derive from the premise that psychopathology consists of maladaptive habits learned in the past. So therapy focuses on extinguishing those habits and reinforcing more adaptive habits. Unlike psychodynamic therapy, this modality has been subject to serious outcome study, and overall it is mildly to moderately effective (Seligman, 2007). Even in its most cutting-edge forms, however, undoing the past is the central task of all the behavior therapies. Exposure therapy (Foa et al., 1999), for example, used to treat post-traumatic stress disorder (PTSD) following rape consists in rehearsing the traumatic experience in the safety of the therapist's consulting room in order to extinguish anxiety and habituate. Whereas some psychopathology such as PTSD is not opaque to the driven-by-the past framework, other psychopathology such as worry—which is focused on an awful future—is entirely opaque.

Cognitive therapy (often called "cognitive-behavior therapy") was spun off from behavior therapy and shares behavior therapy's driven-by-the-past parentage. But it has been relatively unfettered by history, and it is less concerned with the past, or at least the distant past, than is psychodynamic therapy or pure behavior therapy. Human desperation has a way of disregarding theory and devising techniques that relieve suffering. So, for example, the treatment of anxiety is replete with techniques in which scary futures are anticipated (Wolpe, 1958). In problem-solving therapies, clients routinely generate plans of action (e.g., Nezu, 1986). Clients role-play future difficult situations that have not yet occurred (e.g., Corsini, 2010). And many more therapeutic moves that rely on prospection exist. Indeed, there exists a form of cognitive therapy called "future-directed therapy" (Vilhauer et al., 2011).

To illustrate how well a navigating-into-the-future framework fits, consider as examples the following prospective reformulations of several disorders:

- Agoraphobia: the fear that if I go out in public, I will become sick or go crazy.
- Panic disorder: the fear that if my heart starts pounding, I will have a heart attack.
- Obsessive–compulsive disorder: the fear that if I do not flush the toilet in multiples of three, disaster will strike.
- Generalized anxiety disorder: the chronic expectation that something unspecified but awful will soon occur.
- Major depressive disorder: the chronic expectation that the future will be miserable and that if I try to improve things, I will be helpless to make it any better.

In such formulations, what has gone wrong is a maladaptive (and often mistaken) if–then prospection. We do not doubt that the past plays a large role in bringing such beliefs about. We speculate, however, that working directly on the mistaken belief about the future will be at least as effective as revisiting the source of the belief, which is often inaccessibly buried under the detritus of the past. There are at least five ways a therapist can assist the patient by dealing explicitly with maladaptive prospection.

1. Enhancing the prospection of alternatives.
2. Developing more effective prospection.
3. Disconfirming unrealistic prospections.
4. Incentivizing the future.
5. Building meaning and purpose.

Enhancing the prospection of alternatives. In the cognitive therapy of panic disorder, for example, a patient might believe that when she feels her heart pounding, she will go on to have a heart attack, and this belief leads to a spiral of increasing panic. The therapist poses the possibility that a pounding heart will not lead to a heart attack. She suggests an alternative future, previously unconsidered, that a pounding heart is just a normal symptom of mounting anxiety and that knowing this will itself stop the upward cascade of anxiety (Clark, 1986). In depression, for example, the only prospection the patient might entertain is "no matter what I try, I will still not get into graduate school." The therapist might suggest broadening of prospection by working through an entire range of other possible routes to graduate school ("volunteering for work in the professor's lab" or "taking a summer course in advanced statistics") or alternatives to going to graduate school ("Teach for America").

Increasing a patient's ability to generate appropriate alterative prospections on her own will involve increasing her affective as well as imaginative skills. If we are right that perspective taking and empathetic simulation play a vital role in effective prospection, then the importance of both kinds of skills becomes yet clearer. Future-looking therapies therefore would devote some of the time now dedicated to the exploration of past events instead to helping the patient explore possible future situations. Training in empathy is now commonly part of medical school curricula, and a recent study found that such training could have a noticeable effect on clinical practice and physician attitudes if it involved behavioral and affective as well as cognitive elements (Jenkins & Fallowfield, 2002). Something similar seems worth trying in a therapeutic setting.

Developing more effective prospection. The translation from settling on a favored action is often prompt and seamless, but failures can occur, especially when the goal—however highly valued—is distant and indistinct while temptation and distraction are proximate and concrete. As the sad history of dieting shows, simply increasing emotional investment in the goal may not solve the problem, nor will ever-deeper probing into individual history. Prospective-oriented therapy can, however, teach individuals strategies that increase their chance of success. For example, Peter Gollwitzer and colleagues have shown that working with the if–then structure of the mind through the formation of situationally cued "subplan implementation intentions" can help individuals achieve better self-regulation and greater success in achieving long-term ends (Gollwitzer & Sheeran, 2006).

Disconfirming unrealistic prospections. In agoraphobia, for example, the prospection that I will inevitably get sick, throw up, and make a fool of myself if I go out in public is a false and unrealistic distortion. The cognitive (or behavior) therapist can disconfirm this by going with the patient on gradually increasingly difficult trips, culminating in a visit to a shopping mall, showing the patient that it will not lead to panic, vomiting, and making a fool of herself (Marks, 1987). In parallel for depression, the prospection that no matter what I do, my boss will despise me is often an unrealistic distortion. Proposing new ways of doing things (handing in reports one week early, being in the office before the boss arrives, and staying until after she leaves) will disconfirm the distortion of helplessness. The heart of this tack is evaluating whether the prospections that the patient has are unrealistic distortions and then making plans for achieving better futures while disconfirming the distortions.

Incentivizing the future. Future-oriented treatments have often been surprisingly effective. Volpp et al. (2009) approached all smokers employed by a large corporation (including many who had no particular intention of quitting) to take part in a study. All were given information about the benefits associated with quitting and about locally available resources to facilitate quitting. By random assignment, some were also offered a program of financial incentives for quitting, which would pay up to $750 for successful abstinence over a year and a half (to be verified by biochemical test). In principle, long-term financial incentives should make no difference to someone in the grip of a compulsive addiction. Yet the incentive tripled the quit rate. Remarkably, over a dozen participants quit for a year and a half so as to earn the full reward and then resumed smoking. Using incentives to enhance the value of difficult-to-achieve futures, especially if done in futures that are easy to call vividly to mind, is an important complement to providing better information about such futures.

Even apart from therapy, incentives are often influential. The theory of addiction as involuntary or compulsive behavior implies that addicts should be essentially indifferent to changes in price of their preferred substance, but the facts contradict that theory. Tobacco smokers cut back when the price of cigarettes is raised, a fact that has remained true over time (Burns & Warner, 2003). In fact, heavy smokers seem more responsive than light smokers to changes in price, contrary to the view that heavy smokers have no voluntary control over smoking. Even heroin users respond to changes in price, and if they are asked to choose between heroin and money, their decisions depend on how much money is involved (Hart & Krauss, 2008).

Building meaning and purpose. There is growing evidence that a strong sense of meaning and purpose—which we regard as a paradigm instance of robust future orientation—is highly protective against psychopathology (Damon, 2008). In one dramatic example, 84 soldiers who committed suicide had all taken the same test of strengths and weaknesses months before; those soldiers in the very lowest percentage of meaning (strongly disagreeing with "my life has meaning") were at extreme risk for suicide (Lester, Harms, Bulling, Herian, & Spain, 2011). This suggests that building a foundation of meaning and purpose in life should be a major focus of therapy. The lifetime of work by Brian Little (1996) also exemplifies this prospective approach: Little had his subjects identify their core sustainable personal projects—for example, becoming a physician, making oneself more lovable, helping to solve world hunger—and he argued that identity is centered on such prospection. Much of our own thinking about therapy emerges from work on hope and optimism (Seligman, 1990; Snyder, 2002), which then finds its way into positive psychology and prevention (Seligman, 2011). In this endeavor, schoolchildren (positive education) and soldiers (comprehensive soldier fitness) are taught a set of skills in a preventive mode. Three sets of skills are taught: "mental toughness," "identifying and using signature strengths," and building new "social skills," such as active-constructive responding to good events. Given that these are preventive skills, all instruction focuses on how to use these skills in the future. The use of the past in these preventive modalities is largely to illustrate previous failures, which make the need for new skills poignant.

We reiterate that the field of psychotherapy has not been wholly impeded by the driven-by-the past framework that underpins it historically. Our aim in this section was not to dismiss these important developments in therapy, nor to review them, but rather to flesh out a coherent framework that makes more sense of them than does the driven-by-the past framework. From a therapeutic standpoint, however, perhaps the most problematic aspect of the view that we are driven by the past is that it transformed the discipline of psychology into the modern equivalent of predestination. In our view, this framework warped therapeutic theory and practice, by focusing scrutiny on the individual's past, which now lies beyond her control, and turning attention away from the future and the ways in which it will depend on the choices she can make. It should be clear from the examples above that psychotherapy has broken out of this prison and has created a panoply of effective techniques that are best understood in a navigating-the-future framework.

Conclusion

The view that behavior is driven by the past dominated scientific thinking about behavior for many years. Habits and drives do play a role in human life. One need not prospect the future to crank the bicycle pedals, sing "birthday" after "happy," or pocket the keys after locking the door. And experiencing compulsion, addiction, or salt deprivation is much like being driven by the past. But such phenomena are only a fraction of life, and even the compulsive's effort to find a way to wash his hands surreptitiously at a social gathering or an addict's improvisation of a way to get a fix is guided by prospection. The driven-by-the-past framework was psychological Laplacianism, in which habits and drives were claimed to be universally applicable models of learning, memory, decision making, motivation, and cognition. Originally, this framework helped the human sciences escape the idle teleology that the physical sciences cast off so successfully with the Galilean Revolution. The result was a long evolution in thought that culminated in behaviorism and Freudianism. These are now largely behind us, but a commitment to a Laplacian style of explanation continues to have proponents. Productive as this commitment has been, we believe it has outlived its usefulness and that research in a wide range of domains is now pointing toward a new synthesis organized by the idea that intelligent behavior navigates the future (Hawkins & Blakeslee, 2004).

Just as we are convinced that prospecting possible futures is a fundamental explanation of much of human and animal psychology, so we are also impressed by how much there is yet to learn about it.

- Where do the representations of the possible futures we simulate come from? How is past experience used in forming these representations, and what other factors might influence them?
- Beyond food value, monetary value, and risk, which dimensions figure in the evaluation of possible acts and outcomes, and how do they combine to guide selection? What does the decision process look like in detail?
- How is the integration of prospective and habitual control of actions effected, and what role does consciousness play in this process?
- How can prospection and individual differences in prospective abilities be measured validly and used diagnostically?
- What are the cognitive and neural mechanisms of prospection?
- Can prospection and imagination (its first cousin) be improved? Can we teach the next generation of young people to be better simulators, evaluators, and implementers of the future?

In the end, a distinction between what is navigating the future and what is driven by the past is necessary in science. If an ice cube melts, we can explain this entirely in terms of statistical mechanics. To posit a telos of matter seeking an equilibrium temperature would be mere mystification—no information about any possible future state plays a role in the melting process. This is similarly true for reflexive responses, tics, and the most elementary habits. But if a friend helps you to move out of your apartment, very likely your friend's favorable representation of a yet-to-be-realized state did play an essential role. Such explanations are indeed teleological, but that is because nature contains purposeful or goal-oriented organisms, not because nature itself has purposes or goals. So there need be nothing illicit or contrary to the natural order about invoking representations of the future to explain behavior in the here and now. On the contrary, as increasing knowledge of the brain reveals, explanations that leave out this teleological element in the guidance of action are inadequate.

Being driven by the past is as unsuitable as a framework for living as it is for theorizing. It is clear from daily life how much people's evaluations, imaginations, and choices make a difference. Hoping, planning, saving for a rainy day, worrying, striving, voting, risking or minimizing risk, even undertaking therapy, all have in common the presupposition that which future will come about is contingent on our deliberation and action. We have argued that this is no illusion. Prospection is not mysterious, and navigating in light of prospection is at the very core of human action.

A framework is just that—we are not trying to have the last word on the subject or to adjudicate ultimate metaphysics. But after over 150 years of failing to establish that the past drives human action, we suggest that the old, backward-looking framework is no longer productive and that the new, forward-looking framework has much brighter prospects.

Acknowledgments

The authors would like to thank John Anderson, Aaron Beck, Paul Bloom, Kent Berridge, Charles Butter, Phoebe Ellsworth, Rochel Gelman, Daniel Gilbert, Stephen Hinshaw, John Kihlstrom, Ezequiel Morsella, Stephanie Preston, Esa Saarinen, Daniel Schacter, Barry Schwartz, Mandy Seligman, and Bethany Teachman for helpful comments on prior versions of this article. Our gratitude to them does not imply that they agree with the ideas herein.

Declaration of Conflicting Interests

The authors declared that they had no conflicts of interest with respect to their authorship or the publication of this article.

Funding

Support for this article was provided by the Robert Wood Johnson Foundation's Pioneer Portfolio, through a grant, "Exploring Concepts of Positive Health," to Martin Seligman; by the John Marks Templeton Foundation, through a grant, "Positive Neuroscience," to Martin Seligman; and by the National Humanities Center to Peter Railton.

The John Marks Templeton Foundation has launched a major initiative to fund research in all these areas. We would be glad to hear from researchers who would like to contribute to the effort to further characterize and assess prospection and prospective guidance of action.

References

Abegg, M., Manoach, D. S., & Barton, J. J. (2011). Knowing the future: Partial foreknowledge effects on the programming of prosaccades and antisaccades. *Vision Research*, *51*, 215–221.

Addis, D. R., Cheng, T., Roberts, R. P., & Schacter, D. L. (2010). Hippocampal contributions to the episodic simulation of specific and general future events. *Hippocampus*, *21*, 1045–1052.

Addis, D. R., Pan, L., Vu, M. A., Laiser, N., & Schacter, D. L. (2009). Constructive episodic simulation of the future and the past: Distinct subsystems of a core brain network mediate imagining and remembering. *Neuropsychologia*, *47*, 2222–2238.

Ainge, J. A., Tamosiunaite, M., Worgotter, F., & Dudchenko, P. A. (2012). Hippocampal place cells encode intended destination, and not a discriminative stimulus, in a conditional T-maze task. *Hippocampus*, *22*, 534–543.

Andreasen, N. C., O'Leary, D. S., Cizadlo, T., Arndt, S., Rezai, K., Watkins, G. L., . . . Hichwa, R. D. (1995). Remembering the past: Two facets of episodic memory explored with positron emission tomography. *American Journal of Psychiatry*, *152*, 1576–1585.

Anscombe, G. E. M. (1957). *Intention*. Oxford, England: Blackwell.

Aristotle. (1999). *Nicomachean ethics* (T. Irwin, Trans.). Indianapolis, IN: Hackett. (Original work published ca. 330 BC)

Aslin, R. N., Saffran, J. R., & Newport, E. L. (1998). Computation of conditional probability statistics by 8-month-old infants. *Psychological Science*, *9*, 321–324.

Åström, K. J., & Murray, R. M. (2008). *Feedback systems: An introduction for scientists and engineers*. Princeton, NJ: Princeton University Press.

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

Bacon, F. (1959). De argumentis scientarium. As quoted in F. C. Copleston (Ed.), *History of philosophy: Ockam to Suarez* (p. 296). Westminster, MD: Newman Press. (Original work published 1623)

Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review*, *84*, 191–215.

Bargh, J. A., & Chartrand, T. L. (1999). The unbearable automaticity of being. *American Psychologist*, *54*, 462–479.

Bargh, J. A., & Ferguson, M. J. (2000). Beyond behaviorism: On the automaticity of higher mental processes. *Psychological Bulletin*, *126*, 925–945.

Bartlett, R. (1932). *Remembering*. Cambridge, England: Cambridge University Press.

Baumeister, R. F., Crescioni, A., & Alquist, J. (2011). Free will as advanced action control for human social life and culture. *Neuroethics*, *4*, 1–11.

Baumeister, R. F., & Masicampo, E. J. (2010). Conscious thought is for facilitating social and cultural interactions: How mental simulations serve the animal-culture interface. *Psychological Review*, *117*, 945–971.

Beck, A. T., Freeman, A., & Davis, D. D. (2003). *Cognitive therapy of personality disorders*. New York, NY: Guilford Press.

Beck, A. T., Rush, A. J., Shaw, B. F., & Emery, G. (1979). *Cognitive therapy of depression*. New York, NY: Guilford Press.

Behrens, T. E. J., Woolrich, M. W., Walton, M. E., & Rushworth, M. F. S. (2007). Learning the value of information in an uncertain world. *Nature Neuroscience*, *10*, 1214–1221.

Berridge, K. (2004). Motivation concepts in behavioral neuroscience. *Physiology & Behavior*, *81*, 179–209.

Bissmarck, F., Nakahara, H., Doya, K., & Hikosaka, O. (2008). Combining modalities with different latencies for optimal motor control. *Journal of Cognitive Neuroscience*, *20*, 1966–1979.

Biswal, B., Yetkin, F. Z., Haughton, V. M., & Hyde, J. S. (1995). Functional connectivity in the motor cortex of resting human brain using echo-planar MRI. *Magnetic Resonance Medicine*, *34*, 537–541.

Block, N. (2008). Consciousness and cognitive access. *Proceedings of the Aristotelian Society*, *108*, 289–317.

Brunswik, E. (1951). The probability point of view. In M. H. Marx (Ed.), *Psychological theory*. New York, NY: Macmillan.

Buckner, R. L., & Carroll, D. C. (2007). Self-projection and the brain. *Trends in Cognitive Science*, *11*, 49–57.

Burns, D., & Warner, K. (2003). Smokers who have not quit: Is cessation more difficult and should we change our strategies? In D. Burns (Ed.), *Those who continue to smoke: Is abstinence harder and do we need to change our interventions?* (NIH Publication No. 03-5370, National Cancer Institute, Smoking and Tobacco Control Monograph No. 15, pp. 11–31). Bethesda, MD: USDHHS, PHS, NIH, NCI.

Byrne, R. M. J. (2002). Mental models and counterfactual thoughts about what might have been. *Trends in Cognitive Sciences*, *6*, 426–431.

Carnap, R. (1950). *The logical foundations of probability*. Chicago, IL: University of Chicago Press.

Carver, C. S., & Scheier, M. F. (1990). Origins and functions of positive and negative affect: A control-process view. *Psychological Review*, *97*, 19–35.

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

Christoff, K., Gordon, A. M., Smallwood, J., Smith, R., Schooler, J. W., & Posner, M. I. (2009). Experience sampling during fMRI reveals default network and executive system contributions to mind wandering. *Proceedings of the National Academy of Sciences of the United States of America*, *106*, 8719–8724.

Clark, D. (1986). A cognitive approach to panic. *Behaviour Research and Therapy*, *24*, 461–470.

Colwill, R., & Rescorla, R. (1985). Instrumental responding remains sensitive to reinforcer devaluation after extensive training. *Journal of Experimental Psychology: Animal Behavior Processes*, *11*, 520–536.

Conant, R. C., & Ashby, W. R. (1970). Every good regulator of a system must be a model of that system. *International Journal of Systems Science*, *1*, 89–97.

Corsini, R. (2010). *Role playing in psychotherapy*. Piscataway, NJ: Transaction Press.

Courville, A. C., Daw, N. D., & Touretzky, D. S. (2006). Bayesian theories of conditioning in a changing world. *Trends in Cognitive Sciences*, *10*, 294–300.

Craig, A. D. (2009). How do you feel—Now? The anterior insula and human awareness. *Nature Reviews Neuroscience*, *10*, 59–70.

Damon, W. (2008). *The path to purpose*. New York, NY: Free Press.

Darwin, C. (1956). *Origin of the species*. London, England: J. M. Dent & Sons. (Original work published 1859)

Dayan, P., Sham, K., & Montague, P. R. (2000). Learning and selective attention. *Nature Neuroscience*, *3* (Suppl.), 1218–1223.

Decety, J., & Chaminade, T. (2003). When the self represents the other: A new cognitive neuroscience view on psychological identification. *Consciousness and Cognition*, *12*, 577–596.

Demurie, E., De Corel, M., & Roeyers, H. (2011). Empathic accuracy in adolescents with autism spectrum disorders and adolescents with attention-deficit/hyperactivity disorder. *Research in Autism Spectrum Disorders*, *5*, 126–134.

De Neys, W. (2006). Dual processing in reasoning: Two systems but one reasoner. *Psychological Science*, *17*, 428–433.

Denrell, J. (2007). Adaptive learning and risk taking. *Psychological Review*, *114*, 177–187.

Derdikman, D., & Moser, E. I. (2010). A manifold of spatial maps in the brain. *Trends in Cognitive Sciences*, *14*, 561–569.

DeWall, C. N., Baumeister, R. F., & Masicampo, E. J. (2008). Evidence that logical reasoning depends on conscious processing. *Consciousness and Cognition*, *17*, 628–645.

Diedrichsen, J., Shadmehr, R., & Ivry, R. B. (2009). The coordination of movement: Optimal feedback control and beyond. *Trends in Cognitive Sciences*, *14*, 31–39.

Dollard, J., & Miller, N. E. (1950). *Personality and psychotherapy: An analysis in terms of learning, thinking, and culture*. New York, NY: McGraw-Hill.

Draine, S. C. (1997). *Analytic limitations of unconscious language processing* (Unpublished doctoral dissertation). Seattle, WA: University of Washington.

Dugatkin, L. A. (2004). *Principles of animal behavior*. New York, NY: W. W. Norton.

Earman, J. (1992). *Bayes or bust?* Cambridge, MA: MIT Press.

Erev, I., & Barron, G. (2005). On adaptation, maximization, and reinforcement learning among cognitive strategies. *Psychological Review*, *112*, 912–931.

Eykhoff, P. (1994). Every good regulator of a system must be a model of that system. *Modeling, Identification and Control*, *15*, 135–139.

Foa, E., Dancu, C., Hembree, E., Jaycox, L., Meadows, E., & Street, G. (1999). Comparison of exposure therapy, stress inoculation training, and their combination for reducing posttraumatic stress disorder in female assault victims. *Journal of Consulting and Clinical Psychology*, *67*, 194–200.

Fox, M. C., Ericsson, K. A., & Best, R. (2011). Do procedures for verbal reporting of thinking have to be reactive? A meta-analysis and recommendations for best reporting methods. *Psychological Bulletin*, *137*, 316–344.

Frankfurt, H. (1988). *The importance of what we care about*. Cambridge, England: Cambridge University Press.

Freud, S. (1975). *Beyond the pleasure principle* (J. Strachey, Trans.). New York, NY: W. W. Norton. (Original work published 1920)

Freud, S. (2002). *The psychopathology of everyday life* (A. Bell, Trans.). London, England: Penguin. (Original work published 1901)

Gallagher, H. L., & Frith, C. D. (2003). Functional imaging of "theory of mind." *Trends in Cognitive Science*, *7*, 77–83.

Gallistel, C. R., & Gibbon, J. (2000). Time, rate and conditioning. *Psychological Review, 107*, 289–344.

Gilbert, D. (2006). *Stumbling on happiness*. New York, NY: Knopf.

Gilbert, D., & Wilson, T. (2006). Miswanting: Some problems in the forecasting of future affective states. In S. Lichtenstein & P. Slovic (Eds.), *The construction of preferences* (pp. 550–563). Cambridge, England: Cambridge University Press.

Gilbert, D., & Wilson, T. (2007). Prospection: Experiencing the future. *Science*, *351*, 1351–1354.

Glimcher, P. W. (2011). *Foundations of neuroeconomic analysis*. Oxford, England: Oxford University Press.

Goldman, A. I. (2008). *Simulating minds: The philosophy, psychology, and neuroscience of mindreading* (1st ed.). New York, NY: Oxford University Press.

Gollwitzer, P., & Sheeran, P. (2006). Implementation intentions and goal achievement: A meta-analysis of effects and processes. *Advances in Experimental Social Psychology*, *38*, 69–119.

Good, I. J. (1960). Weight of evidence, corroboration, explanatory power, and the utility of experiments. *Journal of the Royal Statistical Society: Series B*, *2*, 319–331.

Grabenhorst, F., & Rolls, E. T. (2011). Value, pleasure, and choice in the ventral prefrontal cortex. *Trends in Cognitive Sciences*, *15*, 56–67.

Greenwald, A. G., & Liu, T. J. (1985). Limited unconscious processing of meaning. *Bulletin of the Psychonomic Society*, *23*, 292–313.

Greicius, M. D., Krasnow, B., Reiss, A. L., & Menon, V. (2003). Functional connectivity in the resting brain: A network analysis of the default mode hypothesis. *Proceedings of the National Academy of Sciences of the United States of America*, *100*, 253–258.

Gupta, A. S., van der Meer, M. A. A., Touretzky, D. S., & Redish, A. D. (2010). Hippocampal replay is not a simple function of experience. *Neuron*, *65*, 695–705.

Gusnard, D. A., Akbudak, E., Shulman, G. L., & Raichle, M. E. (2001). Medial prefrontal cortex and self-referential mental activity: Relation to a default mode of brain function. *Proceedings of the National Academy of Sciences of the United States of America*, *98*, 4259–4264.

Haidt, J. (2001). The emotional dog and its rational tail. *Psychological Review*, *108*, 814–834.

Hart, C., & Krauss, R. (2008). Human drug addiction is more than faulty decision-making. *Behavioral and Brain Sciences*, *31*, 448–449.

Hassin, R., Uleman, J. S., & Bargh, J. A. (Eds.). (2003). *The new unconscious*. Cambridge, MA: MIT Press.

Hawkins, J., & Blakeslee, S. (2004). *On intelligence.* New York, NY: Henry Holt.

Hobbes, T. (1994). *Leviathan* (E. Curley, Ed.). Indianapolis, IN: Hackett. (Original work published 1651)

Hull, C. L. (1943a). *Principles of behavior*. New York, NY: Appleton-Century.

Hull, C. L. (1943b). The problem of intervening variables in molar behavior theory. *Psychological Review*, *50*, 273–291.

Isoda, M., & Hikosaka, O. (2007). Switching from automatic to controlled action by monkey medial frontal cortex. *Nature Neuroscience*, *10*, 240–248.

James, W. (1890). *The principles of psychology* (2 vols.). Cambridge, MA: Harvard University Press.

Jeffrey, R. C. (1953). *The logic of decision*. Chicago, IL: University of Chicago Press.

Jenkins, V., & Fallowfield, L. (2002). Can communication skills training alter physicians? Beliefs and behavior in clinics. *Journal of Clinical Oncology*, *20*, 765–769.

Ji, D., & Wilson, M. A. (2007). Coordinated memory reply in the visual cortex and hippocampus during sleep. *Nature Neuroscience*, *10*, 100–107.

Johnson, A., & Redish, A. (2007). Neural ensembles at CA3 transiently encode paths forward of the animal at a decision point. *Journal of Neuroscience*, *27*, 12176–12189.

Johnson, A., van der Meer, M., & Redish, A. (2007). Integrating hippocampus and striatum in decision-making. *Current Opinion in Neurobiology*, *17*, 692–697.

Kahneman, D. (2011). *Thinking, fast and slow*. New York, NY: Farrar, Strauss, and Giroux.

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, *47*, 263–291.

Kendler, K. S., Walters, E. E., & Kessler, R. C. (1997). The prediction of length of major depressive episodes: Results from an epidemiological sample of female twins. *Psychological Medicine*, *27*, 107–117.

Kennerley, S. W., Walton, M. E., Behrens, T. E. J., Buckley, M. J., & Rushworth, M. F. S. (2006). Optimal decision making and the anterior cingulate cortex. *Nature Neuroscience*, *7*, 940–947.

Killcross, S., & Coutureau, E. (2003). Coordination of actions and habits in the medial prefrontal cortex of rats. *Cerebral Cortex*, *13*, 400–408.

Knudsen, E. I., Heckman, J. L., Cameron, J. L., & Shonkoff, J. P. (2006). Economic, neurobiological, and behavioral perspectives on building America's future workforce. *Proceedings of the National Academy of Sciences of the United States of America*, *103*, 10155–10162.

Kringelbach, M., & Berridge, K. (2009). Towards a functional neuroanatomy of pleasure and happiness. *Trends in Cognitive Science*, *13*, 479–487.

Laird, A. R., Fox, P. M., Eickhoff, S. B., Turner, J. A., Ray, K. L., McKay, D. R., . . . Fox, P. T. (2011). Behavioral interpretations of intrinsic connectivity networks. *Journal of Cognitive Neuroscience*, *23*, 4022–4037.

Lamm, C., Batson, C. D., & Decety, J. (2007). The neural substrate of human empathy: Effects of perspective-taking and cognitive appraisal. *Journal of Cognitive Neuroscience*, *19*, 42–58.

Land, M., & McLeod, P. (2000). From eye movements to actions: How batsmen hit the ball. *Nature Neuroscience*, *3*, 1340–1345.

Langston, R. F., Ainge, J. A., Couey, J. J., Canto, C. B., Bjerknes, T. L., Witter, M. P., . . . Moser, M.-B. (2010). Development of the spatial representation system in the rat. *Science*, *328*, 1576–1580.

Laplace, P.-S. (1951). *A philosophical essay on probabilities* (F. W. Truscott & F. L. Emory, Trans.). New York, NY: Dover. (Original work published 1814)

Lashley, K. (1923). The behavioristic interpretation of consciousness. I. *Psychological Review*, *30*, 237–272.

Lashley, K. (1929). *Brain mechanisms and intelligence*. Chicago, IL: University of Chicago Press.

Le Mens, G., & Denrell, J. (2011). Rational learning and information sampling: On the "naivety" assumption in sampling explanations of judgment biases. *Psychological Review*, *118*, 379–392.

Lester, P. B., Harms, P. D., Bulling, D. J., Herian, M. N., & Spain, S. M. (2011). *Evaluation of relationships between reported resilience and soldier outcomes: Negative outcomes* (Suicide, drug use, & violent crimes, Report No. 1). Available from www.dtic.mil

Lieberman, M. D., Gaunt, R., Gilbert, D. T., & Trope, Y. (2002). Reflexion and reflection: A social cognitive neuroscience approach to attributional inference. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 34, pp. 199–249). San Diego, CA: Academic Press.

Little, B. R. (1996). Free traits, personal projects and idio-tapes: Three tiers for personality research. *Psychological Inquiry*, *8*, 340–344.

Lorenz, K. (1966). *On aggression* (M. K. Wilson, Trans.). New York, NY: Harcourt, Brace & World.

Mackay, D. G. (1973). Aspects of the theory of comprehension, memory and attention. *The Quarterly Journal of Experimental Psychology*, *25*, 22–40.

March, J. G. (1996). Learning to be risk averse. *Psychological Review*, *103*, 309–319.

Marks, I. M. (1987). *Fears, phobias, and rituals*. New York, NY: Oxford University Press.

Marx, K. (1988). *Economic and philosophic manuscripts of 1844* (M. Milligan, Trans.). New York, NY: Prometheus Books. (Original work published 1844)

Mason, M. F., Norton, M. I., Van Horn, J. D., Wegner, D. M., Grafton, S. T., & Macrae, C. N. (2007). Wandering minds: The default network and stimulus-independent thought. *Science*, *315*, 393–395.

McGuigan, F. J., & Ban, T. (1987). *Critical issues in psychology, psychiatry, and physiology: A memorial to W. Horsley Gantt*. Amsterdam, The Netherlands: Gordon and Breach Science Publishers.

Mele, A. R. (2001). *Autonomous agents: From self-control to autonomy*. New York, NY: Oxford University Press.

Merton, R. (1949). *Social theory and social structure*. New York, NY: Free Press.

Miller, G., Galanter, E., & Pribam, K. (1960). *Plans and the structure of behavior*. New York, NY: Holt, Rhinehart, and Winston.

Miller, N. E., & Kessen, M. L. (1952). Reward effects of food via stomach fistula compared with those of food via mouth. *Journal of Comparative and Physiological Psychology*, *45*, 555–564.

Mischel, W. (1968). *Personality and assessment*. New York, NY: Wiley.

Mischel, W. (1973). Toward a cognitive social learning reconceptualization of personality. *Psychological Review*, *80*, 252–283.

Morsella, E. (2005). The function of phenomenal states: Supramodular interaction theory. *Psychological Review*, *112*, 1000–1021.

Najemnik, J., & Geisler, W. S. (2005). Optimal eye movement strategies in visual search. *Nature*, *434*, 387–391.

Neisser, U. (1966). *Cognitive psychology*. Appleton, WI: Appleton-Century-Crofts.

Neuringer, A. (2002). Operant variability: Evidence, functions, theory. *Psychonomic Bulletin & Review*, *9*, 672–705.

Nezu, A. (1986). Efficacy of a social problem-solving therapy approach for unipolar depression. *Journal of Consulting and Clinical Psychology*, *54*, 196–202.

Northoff, G., Heinzel, A., de Greck, M., Bermpohl, F., Dobrowolny, A., & Panskepp, J. (2006). Self-referential processing in our brain: A meta-analysis of imaging studies on the self. *NeuroImage*, *31*, 450–457.

Ortny, A., Clore, G. L., & Collins, A. (1988). *The cognitive structure of emotions*. New York, NY: Cambridge University Press.

Pessoa, L. (2008). On the relationship between emotion and cognition. *Nature Reviews Neuroscience*, *9*, 148–158.

Postman, L. (1951). Toward a general theory of cognition. In J. H. Rohrer & M. Sherif (Eds.), *Social psychology at the crossroads* (pp. 242–272). New York, NY: Harper.

Preuschoff, K., Bossaerts, P., & Quartz, S. (2006). Neural differentiation of expected reward and risk in human subcortical structures. *Neuron*, *51*, 381–390.

Quartz, S. R. (2009). Reason, emotion, and decision-making: Risk and reward computation with feeling. *Trends in Cognitive Sciences*, *13*, 209–215.

Raichle, M. E., Macleod, A. M., Snyder, A. Z., Powers, W. J., Gusnard, D. A., & Shulman, G. L. (2001). A default mode of brain function. *Proceedings of the National Academy of Sciences of the United States of America*, *98*, 676–682.

Railton, P. (2002). Kant meets Aristotle where reason meets appetite. In C. U. Moulines & K.-G. Niebergall (Eds.), *Argument und analyse* (pp. 275–293). Paderborn, Germany: Mentis.

Railton, P. (in press). The affective dog and its rational tale. *Ethics*.

Rangel, A., Camerer, C., & Montague, P. R. (2008). A framework for studying the neurobiology of value-based decision making. *Nature Reviews Neuroscience*, *9*, 545–556.

Rescorla, R. A. (1968). Probability of shock in the presence and absence of CS in fear conditioning. *Journal of Comparative and Physiological Psychology*, *66*, 1–5.

Rescorla, R. A. (1988). Pavlovian conditioning: It's not what you think it is. *American Psychologist*, *43*, 151–160.

Rescorla, R. A., & Solomon, R. L. (1967). Two-process learning theory: Relationships between Pavlovian conditioning and instrumental learning. *Psychological Review*, *74*, 151–182.

Reynolds, S. M., & Berridge, K. C. (2008). Emotional environments retune the valence of appetitive versus fearful functions in nucleus accumbens. *Nature Neuroscience*, *11*, 423–425.

Roberts, W. A. (2002). Are animals stuck in time? *Psychological Bulletin*, *128*, 473–489.

Robinson, T. E., & Berridge, K. C. (2000). The psychology and neurobiology of addiction: An incentive-sensitization view. *Addiction*, *95*(Suppl. 2), S91–S117.

Rolls, E. T., Tromans, J., & Stringer, S. M. (2008). Spatial scene representations formed by self-organizing learning in a hippocampal extension of the ventral visual system. *European Journal of Neuroscience*, *28*, 2116–2127.

Rosenthal, R., & Jacobson, L. (1968). *Pygmalion in the classroom*. New York, NY: Holt.

Rotter, J. B., Fitzgerald, B. J., & Joyce, J. N. (1954). A comparison of some objective measures of expectancy. *Journal of Abnormal and Social Psychology*, *59*, 111–114.

Rushworth, M. F. S., Walton, M. E., Kennerley, S. W., & Bannerman, D. M. (2004). Action sets and decisions in the medial frontal cortex. *Trends in Cognitive Sciences*, *8*, 410–417.

Sanger, T. D. (1989). Optimal unsupervised learning in a single-layer linear feedforward neural network. *Neural Networks*, *2*, 459–473.

Schacter, D. L., & Addis, D. R. (2007a). The cognitive neuroscience of constructive memory: Remembering the past and imagining the future. *Philosophical Transactions of the Royal Society B*, *362*, 733–786.

Schacter, D. L., & Addis, D. R. (2007b). The ghosts of past and future: A memory that works by piecing together bits of the past may be better suited to simulating future events than one that is a store of perfect records. *Nature*, *445*, 27.

Schacter, D. L., Addis, D. R., & Buckner, R. L. (2008). Episodic simulation of future events: Concepts, data, and applications. *Annals of the New York Academy of Sciences*, *1124*, 39–60.

Schneider, W., & Chein, J. M. (2003). Controlled and automatic processing: Behavior, theory, and biological mechanisms. *Cognitive Science*, *27*, 525–559.

Schultz, W. (2002). Getting formal with dopamine and reward. *Neuron*, *36*, 241–263.

Schwartz, B., Schuldenfrei, R., & Lacey, H. (1978). Operant psychology as factory psychology. *Behaviorism*, *6*, 229–254.

Schwartz, N. (2002). Situated cognition and the wisdom of feelings: Cognitive tuning. In L. F. Barrett & P. Salovey (Eds.), *The wisdom of feeling* (pp. 144–166). New York, NY: Guilford Press.

Schwartz, N., & Clore, G. L. (2003). Mood as information: 20 years later. *Psychological Inquiry*, *14*, 296–303.

Scott, S. H. (2004). Optimal feedback control and the neural basis of volitional motor control. *Nature Reviews Neuroscience*, *5*, 534–546.

Seligman, M. E. P. (1990). *Learned optimism*. New York, NY: Knopf.

Seligman, M. E. P. (2007). *What you can change and what you can't* (2nd ed.). New York, NY: Vintage.

Seligman, M. E. P. (2011). *Flourish*. New York, NY: Free Press.

Seligman, M. E. P., & Johnston, J. C. (1973). A cognitive theory of avoidance learning. In F. J. McGuigan & D. B. Lumsden (Eds.), *Contemporary approaches to conditioning and learning* (pp. 69–110). Washington, DC: Winston and Sons.

Shehzad, Z., Kelly, A. M., Reiss, P. T., Gee, D. G., Gotimer, K., Uddin, L. Q., . . . Milham, M. P. (2009). The resting brain: Unconstrained yet reliable. *Cerebral Cortex*, *19*, 2209–2229.

Shevrin, H., Bond, J. A., Brakel, L. A. W., Hendel, R. K., & Williams, W. J. (1996). *Conscious and unconscious processes: Psychodynamic, cognitive, and neurophysiological convergences*. New York, NY: Guilford Press.

Shulman, G. L., Fiez, J. A., Corbetta, M., Buckner, R. L., Miezin, F. M., Raichle, M. E., & Petersen, S. E. (1997). Common blood flow changes across visual tasks: II. Decreases in cerebral cortex. *Journal of Cognitive Neuroscience*, *9*, 648–663.

Singer, T., Critchley, H. D., & Preuschoff, K. (2009). A common role of insula in feelings, empathy and uncertainty. *Trends in Cognitive Science*, *13*, 334–340.

Skyrms, B. (1996). *Evolution of the social contract*. Cambridge, England: Cambridge University Press.

Skyrms, B. (2004). *The stag hunt and the evolution of social structure*. Cambridge, England: Cambridge University Press.

Smallwood, J., Nind, L., & O'Connor, R. C. (2009). When is your head at? An exploration of the factors associated with the temporal focus of the wandering mind. *Consciousness and Cognition*, *18*, 118–125.

Smallwood, J., Schooler, J. W., Turk, D. J., Cunningham, S. J., Burns, P., & Macrae, C. N. (in press). Self-reflection and the temporal focus of the wandering mind. *Consciousness and Cognition*.

Smith, S. M., Fox, P. T., Miller, K. L., Glahn, D. C., Fox, P. M., Mackay, C. E., . . . Beckmann, C. F. (2009). Correspondence of the brain's functional architecture during activation and rest. *Proceedings of the National Academy of Sciences of the United States of America*, *106*, 13040–13045.

Snyder, C. R. (2002). Hope theory: Rainbows in the mind. *Psychological Inquiry*, *13*, 249–275.

Soetedjo, R., Koyima, Y., & Fuchs, A. F. (2008). Complex spike activity in the oculomotor vermis of the cerebellum: A vectoral error signal for saccade motor learning? *Journal of Neurophysiology*, *100*, 1949–1966.

Spreng, R. N., Mar, R. A., & Kim, A. S. (2009). The common neural basis of autobiographical memory, prospection, navigation, theory of mind and the default mode: A quantitative meta-analysis. *Journal of Cognitive Neuroscience*, *21*, 489–510.

Stillman, T. F., Baumeister, R. F., & Mele, A. R. (2011). Free will in everyday life: Autobiographical accounts of free and unfree actions. *Philosophical Psychology*, *24*, 381–394.

Szpunar, K. K. (2010). Episodic future thought. *Perspectives on Psychological Science*, *5*, 142–162.

Szpunar, K. K., Chan, J. C. K., & McDermott, K. B. (2009). Contextual processing in episodic future thought. *Cerebral Cortex*, *19*, 1539–1548.

Tindell, A. J., Smith, K. S., Berridge, K. C., & Aldridge, J. W. (2010). Dynamic computation of incentive salience: "Wanting" that was never "liked." *Journal of Neuroscience*, *20*, 12220–12228.

Tobler, D. N., Dougherty, J. P., Dolan, R. J., & Schultz, W. (2006). Reward value coding distinct from risk attitude-related uncertainty coding in human reward systems. *Journal of Neurophysiology*, *97*, 1621–1632.

Todorov, E., & Jordan, M. I. (2002). Optimal feedback control as a theory of motor coordination. *Nature Neuroscience*, *5*, 1226–1235.

Tolman, E. C. (1948). Cognitive maps in rats and men. *Psychological Review*, *55*, 189–208.

Tomasello, M. (1999). *The cultural origins of human cognition*. Cambridge, MA: Harvard University Press.

Tulving, E. (2002). Episodic memory: From mind to brain. *Annual Review of Psychology*, *53*, 1–25.

Turner, L. H., Solomon, R. L., Stellar, E., & Wampler, S. N. (1975). Humoral factors controlling food intake in dogs. *Acta Neurobiologiae Experimentalis*, *35*, 491–498.

van den Heuvel, M. P., Mandl, R. C. W., Kahn, R. S., & Hulshoff Pol, H. E. (2009). Functionally linked resting-state networks reflect the underlying structural connectivity architecture of the human brain. *Human Brain Mapping*, *30*, 3127–3141.

Van Hoeck, N., Ma, N., Ampe, L., Baetens, K., Vandekerckhove, M., & Van Overwalle, F. (2012). Counterfactual thinking: An fMRI study on changing the past for a better future. *Social Cognitive and Affective Neuroscience*. Advance online publication. doi:10.1093/scan/nss031

Vilhauer, J. S., Young, S., Kealoha, C., Borrman, J., Ishak, W. W., Rapaport, M. H., . . . Mirocha, J. (2011). Treating major depression by creating positive expectations for the future: A pilot study for the effectiveness of future-directed therapy (FDT) on symptom severity and quality of life. *CNS Neuroscience & Therapeutics*, *18*, 102–109.

Volpp, K., Troxel, A., Pauly, M., Glick, H. A., Puig, A., Asch, D. A., . . . Andrain-McGovern, J. (2009). A randomized, controlled trial of financial incentives for smoking cessation. *New England Journal of Medicine*, *360*, 699–709.

Wicker, B., Keysers, C., Plailly, J., Royet, J.-P., & Gallese, V. (2003). Both of us disgusted in my insula: The common neural basis of seeing and feeling disgust. *Neuron*, *40*, 655–664.

Williamson, T. (2008). *The philosophy of philosophy* (1st ed.). New York, NY: Wiley-Blackwell.

Wolfram, S. (2002). *A new kind of science*. Champaign, IL: Wolfram Media.

Wolpe, J. (1958). *Psychotherapy by reciprocal inhibition*. Stanford, CA: Stanford University Press.

Yarrow, K., Brown, P., & Krakauer, J. W. (2009). Inside the brain of an elite athlete: The neural processes that support high achievement in sports. *Nature Reviews Neuroscience*, *10*, 585–596.

Yin, H. H., Knowlton, B. J., & Balleine, B. W. (2004). Lesions of dorsolateral striatum preserve outcome expectancy but disrupt habit formation in instrumental learning. *European Review of Neuroscience*, *19*, 181–189.

Zuo, X.-N., Di Martino, A., Kelly, C., Shehzad, Z. E., Gee, D. G., Klein, D. F., . . . Milham, M. P. (2010). The oscillating brain: Complex and reliable. *NeuroImage*, *49*, 1432–1445.