BEHAVIORAL AND BRAIN SCIENCES (2014) **37**, 121–175
doi:10.1017/S0140525X13000290

The Selfish Goal: Autonomously operating motivational structures as the proximate cause of human judgment and behavior

Julie Y. Huang
Rotman School of Management, University of Toronto,
Toronto, ON M5S 3E6, Canada
and
College of Business, State University of New York at Stony Brook,
Stony Brook, NY 11794
julie.huang@rotman.utoronto.ca

John A. Bargh
Department of Psychology, Yale University New Haven, CT 06520
john.bargh@yale.edu
www.yale.edu/acmelab

Abstract: We propose the Selfish Goal model, which holds that a person's behavior is driven by psychological processes called goals that guide his or her behavior, at times in contradictory directions. Goals can operate both consciously and unconsciously, and when activated they can trigger downstream effects on a person's information processing and behavioral possibilities that promote only the attainment of goal end-states (and not necessarily the overall interests of the individual). Hence, goals influence a person as if the goals themselves were selfish and interested only in their own completion. We argue that there is an evolutionary basis to believe that conscious goals evolved from unconscious and selfish forms of pursuit. This theoretical framework predicts the existence of unconscious goal processes capable of guiding behavior in the absence of conscious awareness and control (the *automaticity principle*), the ability of the most motivating or active goal to constrain a person's information processing and behavior toward successful completion of that goal (the *reconfiguration principle*), structural similarities between conscious and unconscious goal pursuit (the *similarity principle*), and goal influences that produce apparent inconsistencies or counterintuitive behaviors in a person's behavior extended over time (the *inconsistency principle*). Thus, we argue that a person's behaviors are indirectly selected at the goal level but expressed (and comprehended) at the individual level.

Keywords: autonomy; goals; inconsistency; individual behavior; motivation; priming; proximal; self; self-interest; unconscious

> *Do I contradict myself?*
> *Very well then I contradict myself,*
> *(I am large, I contain multitudes.)*
> —Walt Whitman, *Song of Myself*
> (from *Leaves of Grass*, 1855/1960)

1. Introduction

Why do people behave in such contradictory ways? Traditional psychological approaches have struggled to account for the problem of cross-situational inconsistency (e.g., Mischel & Shoda 1995). Indeed, although the field has identified numerous factors that potentially influence a person's behavior – situational factors and individual differences, physiological need states and philosophies of life, conscious and unconscious mental processes, group norms and transcendent values – these influences are bound to conflict with one another, producing apparent inconsistencies in one's overt, observable behavior. These inconsistencies must be managed – rationalized and accounted for – in order to maintain the appearance of stability and predictability to others. The trust that others have in us is the most important determinant of their impressions of and liking for us (Fiske et al. 2007), and inconsistencies lead to the impression of being a "phony," the most negatively viewed personality trait of all, out of 555 traits in Anderson's (1968) normative ratings.

Here, we present the Selfish Goal model, which holds that these inconsistencies in judgment and behavior can be meaningfully understood as the output of multiple, and in some cases, competing goal influences. Whether conscious or unconscious, every goal essentially programs particular sets of behaviors to be enacted by the person pursuing that goal; however, a single person cannot physically express all of these goal influences simultaneously. Instead, he or she expresses many of these influences one at a time, thereby generating behaviors that appear inconsistent across extended time periods. In other words, contradictions in individual behavior may occur because a person, as Whitman acutely observed, contains multitudes of influences.

The inspiration for this target article is interdisciplinary, as we hope will be its appeal. Our core insight is derived from a prominent theory in evolutionary biology. In *The Selfish*

0140-525X/14 $40.00

Gene (1976), Richard Dawkins describes how every organism is composed of multiple genes, each of which can be seen as using that organism as a survival machine. Through the blind processes of natural selection, genes exert an influence on their host organism's behavior that maximizes their chances of propagation into future generations, sometimes to the detriment of the host organism's life. We believe it is useful to describe a person's psychological structure in an analogous fashion. Thus, we conceptualize each individual as comprised of multiple, oftentimes conflicting *goals*, each of which influences that person in a systematic, "selfish" manner.

As we will argue, human goal pursuit – whether operating consciously or unconsciously – constrains a person's information processing and behaviors in order to increase the likelihood that he or she will successfully attain that goal's end-state. These multiple, sometimes conflicting goals can produce different behaviors, judgments, and even self-representations in the same person that may appear inconsistent or contradictory across time, because they will vary as a function of which goal happens to be most active and motivating in the particular situation. Put another way, observed incoherencies in a person's actions may result because behavior is being selected (and is coherent) at a lower, less apparent goal level.

First, we trace the evolution of adaptive human behavior back to its unconscious and goal-driven roots. We survey evidence that an unconscious, less integrated system coordinated organism-level behavior prior to the evolution of conscious processing capabilities. When consciousness did emerge, we suggest that it was built upon preexisting unconscious processes, thereby producing the striking structural and phenomenal similarities observed between conscious and unconscious goal pursuits in recent psychological research.

Second, we generate predictions from this framework regarding the qualities and outcomes of human goal-driven behavior in the present day. If, as we assume, individual behavior arose from a system of unconscious and selfish processes that predated the evolution of consciousness, we should find contemporary signatures of the unconscious processes operating to guide behavior and judgment independently of conscious guidance (*automaticity*). To the extent that these unconscious systems evolved to facilitate adaptive responses to the environment, and were the proximate control system over those responses, they should be found to produce a "full-system" orientation toward specific end-states – adjusting perceptual, evaluative, cognitive, as well as motivational, parameters (*reconfiguration*) to optimize attainment of the desired end-state.

The present model also predicts that the evolution of consciousness from unconscious processes should be revealed in the observation of highly similar properties between the two denominations of goal pursuit (*similarity*). In addition to the straightforward prediction that unconscious goal pursuit should be characterized by similar processes and outcomes as revealed by a century of research on conscious goal pursuit, The Selfish Goal model makes the nonobvious prediction that conscious goal pursuit also should be found to operate in ways only recently discovered to be true of unconscious goal pursuit. That is, even conscious goals should be found to operate selfishly, even if doing so produces behavioral inconsistencies for the individual (*inconsistency*). Recent findings from social and developmental psychology, as well as neuroscience, will be marshaled in support of each of these principles.

A model such as this, with interdisciplinary theories and implications, is likely to encounter challenges such as establishing shared definitions and drawing from areas of relative inexpertise. Although we acknowledge these challenges, we believe that our theoretical contribution brings important recent developments from social psychology to bear in discussions with other disciplines that share the goal of understanding individual human behavior.

Julie Y. Huang is currently a Postdoctoral Fellow at the Rotman School of Management at the University of Toronto. She received her Ph.D. in psychology at Yale University in 2011. Starting in 2014, she will be an Assistant Professor in the College of Business at Stony Brook University.

John A. Bargh is a Professor of Psychology and Cognitive Science at Yale University. He graduated from the University of Illinois in 1977 and received his Ph.D. in Psychology from the University of Michigan in 1981. Bargh was on the faculty of New York University from 1981 to 2003, when he moved to Yale. Bargh received the Dissertation Award from the Society for Experimental Social Psychology in 1982; the Early Career Contribution Award from the American Psychological Association in 1989; and the Donald Campbell Award for career contribution to social psychology from the Society for Personality and Social Psychology in 2006. In 2011, Bargh was elected to the American Academy for Arts and Sciences.

2. Who – or what – is in control of an individual's behavior?

We all share the intuition that the form and content of our behaviors are produced through conscious, intentional choice and internal processes of which we are aware and able to report on reasonably accurately to others (Wegner 2002). In the past, traditional psychological approaches to human motivation similarly assumed an agentic, conscious self at the helm, deliberately forming judgments, making decisions about which courses of action to take, and then guiding one's behavior along those intentional lines (e.g., Ajzen & Fishbein 1980; Bandura 1986; Baumeister et al. 1998; Locke & Latham 1990; Mischel 1973). In some prominent models, conscious choice of behaviors or goals to pursue was conceptualized as a bottleneck – nothing happened without one's awareness and conscious consent (e.g., Bandura 1986; Locke & Latham 2002, p. 705).

Since then, three modern lines of research have cast doubt on the basic assumptions of conscious-centric control models. First, studies increasingly highlight the power of situational variables in determining behavior, including external influences that override internal sources of control such as self-values and personality (e.g., Darley & Latane 1968; Milgram 1963; Mischel 1973; Ross & Nisbett 1991). Second, research on the limits of introspective access demonstrates that people are often unaware of the reasons behind their actions and the actual sources of their evaluations and subjective feelings about the external world (Bar-Anan et al. 2010; Nisbett & Wilson 1977; Wilson & Brekke 1994) – access they would be expected to have if they were consciously aware of making those choices and deciding what to do. Third, "dual-process" models (e.g., Chaiken & Trope 1999; Morewedge & Kahneman 2010; Posner & Snyder 1975; Strack & Deutsch, in press) hold that external situational influences often operate in an automatic and implicit fashion to directly instigate the higher mental processes involved in information processing and behavior, thereby bypassing the consciousness bottleneck and eliminating the need for an agentic "self" in the selection of all behavioral and judgmental responses (Bargh 2007; Bargh et al. 2012).

Research involving the passive activation or *priming* of higher-order concepts by contextual features combines the situational and automatic emphases of contemporary research. Priming studies have consistently demonstrated that the mere exposure to environmental events is sufficient to directly trigger higher mental processes, in the absence of any conscious intentions or awareness that they operate (see Dijksterhuis et al. 2007; Higgins 1996). Unconscious processes have been shown to produce evaluation and social judgment (Fazio 1986; Ferguson 2008), stereotyping and prejudice (Devine 1989), social behavior (Bargh et al. 1996; Dijksterhuis & van Knippenberg 1998; Loersch & Payne 2012), and goal pursuit (Chartrand & Bargh 1996; Custers & Aarts 2005; Marien et al. 2012), with research participants unaware of the role played by these external stimuli in their decisions and behaviors.

Although important strides have been made, key issues remain. The environment is a rich source of multiple, oftentimes simultaneous cues – many of which are linked to different, often competing behavioral impulses. How are these *parallel* environmental influences funneled into the necessarily *serial* behavior of a single individual (the "reduction problem": Bargh 2006; Merker 2007)? An important aim of the present model is to address this question – namely, to argue that the currently active goal is the proximate controller of present behavior and judgment, overriding other conflicting influences in a "single-minded" fashion if necessary to facilitate goal completion.

2.1. The Selfish Goal

To illustrate how behavior selection might occur at a less-centralized level than that of the individual person (where it is overtly expressed), we describe how goals direct and constrain a person's behavioral possibilities *as if* the goals themselves were selfish.

The Selfish Goal model offers general predictions regarding individual behavior in the current day. The automaticity principle proposes the existence of unconscious processes tied to individual judgment and behavior. As we will review, research has indeed discovered psychological processes that are dissociated from individual awareness or guidance, yet can change how that person sees the world and behaves in response (Bargh et al. 2012). For example, priming (passively and temporarily activating) an individual's internal goal representation affects subsequent judgments and behaviors in a manner consistent with him or her being in a motivated state (Bargh et al. 2001; Bargh et al. 2010; Dijksterhuis & Aarts 2011).

The second principle, reconfiguration, holds that changes in a person's judgments and behaviors during goal operation occur in order to optimize that person's chances of completing the goal. In particular, we focus on experiments featuring unconscious forms of pursuit, because by definition participants are unaware of the goal and hence cannot guide their own perceptions and behavior in a goal-congruent way. Nevertheless, research suggests that people who are unaware that they are pursuing a goal respond to the world in a way that maximizes the likelihood of goal completion, such as by paying more attention to objects in the environment that would assist with goal pursuit and becoming predisposed to like and physically approach those objects. Goals operate autonomously (i.e., independent of guidance from the conscious individual) through these mechanisms to encourage achievement of their associated end-states.

Because all goals are selfish, the similarity principle holds that conscious goal pursuit should resemble its unconscious counterpart in these regards. Recent research reveals a number of striking similarities between conscious and unconscious goal operation, from shared neural correlates and subprocesses to similar autonomous effects upon individual judgment and behavior. Even when a person consciously engages in goal pursuit, one is not necessarily controlling (or even aware of) how that goal has transformed one's experience of the world. As an analogy, when conversing with another person, we are not aware of how our ideas and thoughts are being transformed into grammatical and coherent sequences of words.

The inconsistency principle proposes that as multiple goals within a single individual become active, operate, and turn off, the person pursuing those goals may appear to be acting inconsistently, or in a manner that seems contrary to his or her interests. This is particularly evident at the stage of goal completion, where the mental representation of a goal becomes less accessible but, ironically, the pursuer becomes more likely than usual to exhibit behaviors that contradict the recently completed goal. We will cast these lines of research in the light of evolutionary theories to argue that the proximate control of individual behavior within a situation rests with the currently active goal, and if more than one goal is currently active, then with the most incentivized (important) of these.

2.2. Terminology

Before proceeding, it is important to make clear what we mean by some commonly used terms and concepts. A *goal* is a construct of central importance in psychology, yet little consensus exists regarding its precise definition (Austin & Vancouver 1996; Elliot & Fryer 2008). For example, goals are sometimes equated with related concepts such as needs, motives, biological goals, and drives (e.g., Austin & Vancouver 1996; Pervin 1989;); others draw distinctions (e.g., Gollwitzer & Moskowitz 1996). Some perspectives conceptualize goals as explicit standards for behavior set and regulated by the individual (e.g., Bandura 1986; Locke & Latham 1990); others suggest that they can be triggered by environmental contexts and operate independently of individual awareness through a combination of cognitive and affective mechanisms (e.g., Bargh et al. 2001; Custers & Aarts 2010; Kruglanski et al. 2002).

Here, we follow the standard social-cognitive definition of goals as *mental representations of desired end-states* (Aarts & Dijksterhuis 2000; Bargh 1990; Fishbach & Ferguson 2007; Kruglanski et al. 2002). We are less interested in long-term, chronic goals (e.g., raising children; sometimes referred to as "life tasks"; Cantor & Blanton 1996; Pervin 1989) and more interested in time-limited instrumental behaviors enacted in the current situation (e.g., being helpful to others; satisfying one's hunger). Our focus is less on the very concrete actions that can be described in fully objective terms (e.g., pressing a button; opening a door) and more interested in higher-level end-states that provide those actions with meaning and that come in both conscious and unconscious denominations.[1] A person can reach the same end-state by "operating" on the environment in different ways (e.g., one can be helpful by taking out the trash or by uttering supportive statements); the notion that Skinner (1953, p. 65) captured with the concept of an operant, or a "set of acts" defined by their effect, Kruglanski et al. (2002; also Austin & Vancouver 1996) termed *equifinality* and Lashley (1942) referred to as *motor equivalence,* in which a motoric action goal can be realized in a variety of ways.

Historically, there has been high consensus regarding the features of *conscious* information processing: a person is aware of those processes and intends them to occur; they operate in a serial fashion, are somewhat resource limited at any given time, and are controllable (e.g., Bargh 1984; 1994; Posner & Snyder 1975). In contrast, views of the unconscious have arisen out of the study of processes that do not possess all of these features (see Bargh & Chartrand, 2000). For example, "preconscious" or "preattentive" processes operate automatically upon encountering stimuli in the environment and feed subsequent conscious processes (Neisser 1967; Posner & Snyder 1975), whereas other unconscious processes become automatic through a process of skill acquisition (e.g., Shiffrin & Schneider 1977) *given* the intent to engage those processes in the first place, such as when typing or driving.

By *unconscious,* we refer to information-processing events in the human nervous system that, although capable of influencing a person's behaviors, emotions, cognitions, and motivations, "do not influence subjective experience in a way that [he or she] can directly detect, understand, or report the occurrence or nature of these events" (Godwin et al. 2013; Morsella & Bargh 2011; Nisbett & Wilson 1977). Note that by this definition, one can be consciously aware of a stimulus event but unconscious of its influence on oneself (see Bargh 1992).[2] A process is described here as unconscious even if a person is able to report the presence of a stimulus and respond to it (e.g., in an experiment, one is judging stockings presented in an array, deems the right-most choice of highest quality, and provides reasons why) but lacks awareness regarding the causes and processes underlying the effect (e.g., that one may have been biased by a position effect on evaluation that favored the right-most stocking independent of quality-related factors; Nisbett & Wilson 1977).

This use of the term "unconscious" highlights the *unintentional* nature of the process, along with a lack of awareness of its underlying causes and processes (see Bargh 1992; Bargh et al. 2001). This definition is consonant with uses of the term "unconscious" in both evolutionary theory and historically by psychologists. For example, Darwin (1859), one of the earliest scientists to refer to unconscious processes, employed the term when describing how the farmers and stock breeders of his time made implicit use, without explicit awareness, of the laws of natural selection in order to produce larger ears of corn and fatter sheep. Similarly, Dawkins (1976) wrote of nature as the "blind watchmaker, the unconscious watchmaker" to signify that there was no agentic, intentional guiding hand in producing these adaptive designs (see also Bargh & Morsella 2008; Buss et al. 1998; Dennett 1991; 1995). According to Brill (1938), Sigmund Freud credited Mesmer and the early hypnotists with the original discovery of the unconscious when they caused their subjects, through post-hypnotic suggestion, to behave in ways that they did not intend.

3. The primacy of the unconscious

Why individuals behave as they do within given situations is a fundamental question of social psychology (Aronson 1995), and social-psychological answers to this question should, ideally, be in harmony with relevant knowledge from the natural sciences (Pinker 1997). Accordingly, we describe literature on consciousness and evolutionary biology that provide the theoretical backbone of the Selfish Goal. The following section is not intended as a comprehensive review of either topic, which would be outside of the scope of the present article; rather, areas of consensus are highlighted because they provide central guiding insights.

3.1. Prior to the evolution of consciousness

Understanding how consciousness evolved remains one of the greatest scientific mysteries (Kennedy 2005). For certain portions of the puzzle, however, consensus is beginning to emerge (for review, see Godwin et al. 2013), and these areas of agreement offer insight regarding present-day individual behavior. For example, few experts would object to the claim that, at some point in the evolutionary past, something unconscious became something that was conscious (e.g., Corballis 2007; Deacon 1997; Dennett 1991; Donald 1991; Godwin et al. 2013; Macphail 1998). Much of the activity in the brain and in the rest of the nervous system operates unconsciously, reinforcing the argument that consciousness is not an inherent property of human cognition and likely emerged from an evolutionary landscape that had until then been the province of "not-conscious," or unconscious processes.

Moreover, the phenomena linked to consciousness are themselves associated with but a subset of neuroanatomical regions in the central nervous system (e.g., Dehaene & Naccache 2001; Morsella et al. 2009). Identifying which particular regions are responsible for consciousness remains an issue of active debate, as once-prevalent views of the cerebral cortex as the "organ of consciousness" face direct challenge by theories and evidence linking subcortical structures such as the upper brainstem to the instantiation of conscious states (e.g., Godwin et al. 2013; Merker 2007; Penfield & Jasper 1954).

The evidence that only a subset of processes and regions of the brain are associated with consciousness, plus the fact that humans share much of this unconsciously operating nervous system with earlier-evolving members of the animal kingdom (some of whom arguably lack consciousness), leads to the conclusion that conscious processes are a phylogenetically later adaptation of the brain. As Dennett (1991, p. 171) pointed out, "Since there hasn't always been human consciousness, it has to have arisen from prior phenomena that weren't instances of consciousness."

One can infer additional characteristics of the unconscious system guiding behavior. The existing system may have been unconscious and perhaps less centralized in the absence of consciousness, but it was nevertheless capable of guiding an organism's behavior toward adaptive outcomes. Borrowing an example from Merker (2007), there is little reason to believe that the cubomedusa jellyfish, with its noncephalized nervous system, experiences consciousness (i.e., that there is anything it is "like" to be this jellyfish). Yet, despite the fact that it lacks a brain even remotely like a human's, the jellyfish still exhibits adaptive behaviors, including flexible, goal-driven movements in response to environmental stimuli such as prey.

Put another way, although consciousness is traditionally associated with behavior selection and coordination, this does not mean the older, unconscious, and less-centralized system(s) did not also give rise to adaptive, streamlined behavior at the level of the individual organism. (For examples of sophisticated unconscious control and conflict management, see Morsella 2005; Suhler & Churchland 2009.) Indeed, evidence of behavior directed toward adaptive ends is omnipresent in the animal kingdom – even in creatures not assumed to possess consciousness – so much so that it is, in the view of one evolutionary biologist, "perhaps the most characteristic feature of the world of living organisms" (Mayr 1976, p. 389).

Many behaviors that serve the ultimate "end" goal of all adaptations (i.e., differential reproduction) are also goal-driven on a more proximate level. For example, when Dawkins (1976) identified "apparent purposiveness" as an important property of adaptive behavior, he clarified, "By this I do not just mean that it seems to be well calculated to help the animal's genes to survive, although of course it is. I am talking about a closer analogy to human purposeful behavior … [organisms] that behave as if motivated by a purpose" (p. 50). Indeed, evolutionary biologists and psychologists alike link motivations or goals to adaptive outcomes (Campbell 1974; Cosmides & Tooby 2013; Kenrick et al. 2010; Pinker & Bloom 1990; Popper 1972; Symons 1992; Tetlock 2002; Tomasello et al. 2005; Tooby & Cosmides 1992). A bird starting on its migration, a male displaying to a female, a predator stalking its prey, and the prey's subsequent flight – all are acting toward a specific end-state in the local environment (e.g., reaching safety and shelter; obtaining food to eat).

Goal-related constructs are prominently featured in models of adaptive human behavior, further rooting these behavioral patterns deep in the evolutionary past (Bugental 2000; Cosmides & Tooby 2013; Kenrick et al. 2010; Neuberg et al. 2004). Similar to its use in social cognitive literature, these goal constructs involve "if-then" contingencies in which environmental cues trigger psychological and behavioral phenomena that are designed by natural selection *for a purpose* – to facilitate adaptive outcomes. Accordingly, different goals have their own specific evolutionary function (e.g., attaining safety; securing a mating partner) and are associated with different decision rules, neuroanatomical regions, and behavioral responses. An organism that suddenly encounters predators should be most likely to fight, flee, or freeze; that same creature, in the company of potential mates, should be compelled to act differently by approaching and displaying instead.

Although research suggests that adaptive behavior across the animal kingdom can be organized into goal categories, it is less clear as to how multiple behavioral impulses – each designed for specific functions – are funneled into streamlined behavior. Evolutionarily adaptive goals require different, often incompatible organism-level behaviors for its pursuit (e.g., fleeing vs. eating), which must be expressed in serial fashion at critical moments. Kenrick et al. (2010, p. 303) emphasized this tension as a "continual interplay between motivational systems and the perception of affordances (fitness-relevant threats and opportunities) in the immediate environment." What, then, are the characteristics of an evolutionarily ancient system that could integrate multiple, sometimes competing influences into overt behavior expressed at the level of the individual organism – particularly in the absence of overarching conscious processing to integrate and prioritize these goals?

3.2. *The selfish gene and the Selfish Goal*

Understanding the component pieces of the puzzle offers a preliminary step toward an answer, and evolutionary biology provides a useful metaphor with which to describe their operation. As previously stated, in *The Selfish Gene,* Dawkins (1976) described how genes, through the blind process of natural selection, influence the design of their host organism in order to maximize their chances of propagation into future generations. This theory characterizes genes as essentially "selfish" in that their only concern is their own replication, not the welfare of their host organism, except as it might impact on replication. (Note that this selfish influence on one level can also be seen as cooperative at another level, as Dawkins [1976] pointed out, insofar as multiple genes must coexist and be expressed within a single individual.)

We argue that the dynamic between genes and their host organism is analogous to the relationship between goals and the individual pursuing them. Every organism is comprised of multiple genes, all using that organism as their survival machine into the next generation. Similarly, a person comprises many different goals, each of which operates on that individual to produce successful pursuit of a specific end-state.

Supporting the notion that unconsciously operating goal processes are capable of influencing individual-level outcomes, experiments from both evolutionary psychology and social cognition highlight early-stage orienting mechanisms (e.g., selective attention and perception) that serve as "building blocks" for subsequent human behavior (Balcetis & Dunning 2006; Maner et al. 2008; Neuberg et al. 2004). Given that a person is effectively steered toward particular sets of actions (and simultaneously away from others), these processes can be seen as operating in self-interested ways, as a gene encodes organism-level behavior that reliably promotes replication of that gene, and not necessarily the health or well-being of the host organism itself.

3.3. After the evolution of consciousness

Although little consensus exists regarding *how* consciousness evolved, even the basic claim *that* consciousness evolved in a nervous system composed of unconscious, less centralized processes can be informative. Evolutionary theorists note that given the constraints of building costs and materials, evolution behaves as a "tinkerer," modifying existing structures in a gradual, incremental fashion, instead of creating entirely new ones each time from scratch (Allman 2000; Dawkins 1976; Jacob 1977). Solutions created from this process may not be perfect, but nevertheless satisfice (Dennett 1995; Simon 1956).

As the process of natural selection uses old elements and alters them for new functions, traces of the original structure and function often remain in the end-product. Evidence of this co-option process can also be found across the animal kingdom, as when the vibration-sensitive reptilian jaw evolved into mammalian middle-ear bones, fish fins morphed into forelimbs for land-dwelling creatures, and the esophagi of freshwater fish expanded to serve terrestrial lung functions (for more examples, see Jacob 1977; Pinker 1997).

The human nervous system also bears marks of evolutionary tinkering (e.g., Anderson 2010; Buss et al. 1998; Jacob 1977; Morsella 2005), which suggests that conscious processes were likely shaped by the preexisting unconscious behavioral control system (see Bargh & Morsella 2010). One would expect to see, as one does, that conscious states are selectively integrated into the existing unconscious system, while many behavioral processes remain unaffected. Conscious processes should also retain features of the preexisting elements integrated into its design. Consequently, one might also expect conscious processes to function in a similarly "selfish" manner as do their unconscious counterparts, because both operate on the nervous system in the service of adaptive, individual-level behavior.

3.4. Summary

Inherent in any attempt to integrate fields is the risk of drawing too many assumptions. For this reason, the Selfish Goal relies on a limited number of claims about the evolution and function of consciousness: (1) the relatively late evolutionary emergence of consciousness, leading to the further assumption that (2) some other, unconscious, less centralized system(s) must have been guiding serially expressed, adaptive behavior prior to that point in evolutionary history. Finally, we assume a basic principle of evolution through natural selection, that (3) it occurs through gradual changes to the existing machinery (e.g., Allman 2000), and not via dramatic, complete overhauls of that machinery. In other words, less integrated unconscious goal processes lay down the landscape upon which their conscious counterparts operate today.

Given the theoretical evidence to support these assumptions, we surmise that at a hypothetical Time 1, a less centralized system of unconscious, adaptive processes, including goals, drove behavior within the same biological entity. At Time 2, consciousness evolved within the nervous system and was likely shaped in structure and function by its integration with the preexisting unconscious landscape. By theoretical Time 3 (the present day), we see events indicating further development of conscious processes (e.g., expansion of cortical matter; metacognitive abilities; representation of self; and tighter executive control over the initial impulses to action). Nonetheless, some original unconscious processes from Time 1 remain, operating in concert with and independent of later arriving conscious processes.

In the next section, we consider the implications of this model for behavior in current-day contexts. Evidence for four principles will be offered in support of the claim that a person's judgments and behaviors can be meaningfully examined from the perspective of which goals he or she is currently pursuing. In so doing, we hope to clarify how inconsistencies in a person's behavior might result from a behavioral guidance system comprised of selfishly acting goals.

4. The active goal as proximal cause of human behavior

> *My thinking is first and last and always*
> *for the sake of my doing,*
> *and I can only do one thing at a time.*
> —William James (1890/1981)

Figuratively put, the Selfish Goal model posits a system wherein multiple goals engage in the selection of behavior while the individual person is actually "expressing" it. As multiple conscious and unconscious goals operate, each steering the individual toward specific (and oftentimes conflicting) end-states, inconsistencies in individual behaviors will arise. Note that goals often can and do encourage behaviors that are consistent with (or at least not opposed to) other goals' end-states, or one's consciously preferred outcomes (e.g., exercising self-control by declining dessert is consonant with a dieter's self-reported preferences). We focus on telling instances when the two interests are inconsistent or at odds with each other.

In extreme cases, the tension between the behavioral imperatives issued by the currently active goal and the other priorities of the person pursuing that goal (over time and across situations) can produce trade-offs between what is "good" for the goal being pursued versus what is "good" for the individual. This dynamic is most evident in addictions (e.g., Baker et al. 2004) where the addict reports feeling helpless to resist the urge to consume the drug and may engage in self-destructive behavior (as well as behavior that typically contradicts his or her important self-values, such as lying and stealing) in order to acquire the drug. Decades of research on drug abuse substantiate the similarity between addictions and more standard goal pursuits (e.g., Loewenstein 1996), as expressed colloquially by the phrase, "addiction hijacks the motivational system."

The pursuit of everyday goals has "selfish" effects as well and may cause an individual to desire things that he or she might not have wanted if not actively pursuing the goal. For example, young women primed with the mating goal express more positive attitudes toward and stronger intentions to engage in attractiveness-enhancing yet dangerously unhealthy behaviors such as spending time in tanning booths and taking diet pills (Hill & Durante 2011). Those behaviors may facilitate the currently active goal of mating (by increasing one's sex appeal) but operate to the long-term detriment of the individual. Indeed, participants' attitudes toward these behaviors when the mating goal was not currently active were considerably more negative. In a similar vein, males who are unconsciously pursuing status goals express greater hostility to other males (Griskevicius et al. 2009). In this case, increased willingness to aggress furthers the active goal (by physically intimidating one's competition) but ultimately jeopardizes the health of the individual pursuing that goal, both during pursuit and afterwards.

The Selfish Goal perspective is in harmony with other motivational models in social psychology, particularly the Evolved Hierarchy of Needs (Kenrick et al. 2010) and Functionalist Social-mindset theory (Tetlock 2002). We take a complementary perspective to Kenrick et al. (2010) by focusing on the general *structure* of goal influence, as opposed to domain-specific content, and offering more proximate explanations for behavior rather than the ultimate causes of it (see Killeen 2001). Tetlock's (2002) theory posits the existence of evolved complex mindsets associated with different social motives depending on the individual's current situation and contextual role. When perceived as misbehaving according to group norms, an individual adopts a "defense attorney" mindset that focuses on external reasons for the socially negative action in order to maintain or restore one's good standing in the group; however, the same individual faced with another person's misbehavior adopts a very different "prosecutorial" mindset, with the opposing goal of holding the miscreant's feet to the fire in order to uphold those important group norms.

This model is also inspired by extensive research on Goal Systems Theory (GST; Kruglanski et al. 2002). GST holds that goals are mental representations with cognitive and motivational properties and exist within systems of interconnected goals. Within this goal system, limited cognitive resources are distributed in a constant-sum fashion, which can lead to mental resources being pulled toward and away from particular goals, depending on which is active. For example, the activation of a goal to seek entertainment or to eat a yummy calorie-rich food (i.e., temptations) can trigger self-regulatory goals to get some work done or to eat something healthy instead (Fishbach et al. 2003). We similarly highlight conflict among goals. Whereas GST attributes this tension to limited cognitive resources, our evolutionarily inspired model focuses on conflict within an evolved system that must funnel multiple impulses operating in parallel into streamlined behaviors within a single physical body.

4.1. Predictions from the present model

In section 3 we reviewed theories that hold that unconscious, selfish goals lay the evolutionary foundation for the operation of their conscious counterparts. Specifically, the Selfish Goal model holds that multiple, selfish goals indirectly guide the streamlined behaviors expressed by a single individual. This framework predicts the following about individual judgment and behavior in current-day contexts:

1. *The automaticity principle*. Unconscious processes can influence behavior in the absence of individual awareness or guidance.
2. *The reconfiguration principle.* The most motivating or "active" goal should constrain the individual's information processing and behavioral possibilities in a way that encourages achievement of the goal's end-state.
3. *The similarity principle.* Conscious goal pursuit should recruit similar processes and produce similar outcomes as unconscious goals.
4. *The inconsistency principle.* Temporarily active goals can produce outcomes for the individual that appear inconsistent over time or contrary to the individual's interests.

4.2. The automaticity principle

We will first review evidence for the existence and qualities of unconscious processes that are tied to overt human behavior yet dissociated from individual awareness. Dawkins (1976) notes that adaptive behavior in the animal kingdom "works on a time-scale not of months but of seconds and fractions of seconds. Something happens in the world … in milliseconds nervous systems crackle into action, muscles leap, and someone's life is saved – or lost" (p. 55). The past quarter century of psychological research documents similar, lightning-fast psychological mechanisms in humans (e.g., automatic evaluations and preferences, automatic influences of the perceived environment on behavioral responses, and an automatic mode of goal pursuit; Bargh et al. 2012; Bargh & Ferguson 2000; Greenwald & Banaji 1995), which were observed through methodological advances such as sequential priming and the Implicit Association Test. For example, the affective sequential priming task allows researchers to measure the implicit value of a target (and its associated goal) in terms of the extent to which it automatically facilitates response time to positive or negative concepts. Because less time is needed to categorize a target stimulus as positive or negative when it is preceded by a concept of the same valence (but not otherwise semantically related; e.g., Fazio 1986), researchers using this task are able to examine attitude activation and influence occurring outside of the individual's awareness and control.

Automatic evaluations play an important role in guiding both cognition (Ferguson et al. 2005; Niedenthal 1990; Cantor & Blanton 1996) and overt behavior (the latter supporting the idea that these mechanisms may have served adaptive roles in the past; Roe & Simpson 1958). For example, implicit positive evaluations automatically create approach behavioral dispositions (muscular readiness) toward stimuli and negative automatic evaluations to avoidance (withdrawal) behaviors (Chen & Bargh 1999). Indeed, Ferguson (2008) revealed that positive automatic evaluations toward cooperation-related primes predicted participants' actual helping behaviors on a subsequent task (e.g., volunteering time to assist another person). Moreover, this link between automatic evaluation and muscular readiness has recently been successfully exploited in therapeutic techniques for the treatment of addictions, with patients making incidental avoidance arm movements in response to addiction-related stimuli across hundreds of trials, which has the consequence of significantly reducing their cravings and use of the substance (Wiers et al. 2010).

4.2.1. Conscious awareness or intention is not required for goal pursuit. In one of the initial experimental demonstrations of unconscious goal pursuit, Bargh and colleagues (2001; see also Chartrand & Bargh 1996) found that merely exposing participants to words related to cooperation (e.g., *help, assist*) in the guise of an ostensible language test increased cooperative behavior compared to another condition in which participants were explicitly instructed to cooperate. Following the experimental task, participants were explicitly asked the extent to which they had tried to cooperate during the task. In the explicit cooperation condition, these estimates correlated positively with actual cooperative behaviors (indicating some degree of conscious awareness of the reason for their behavior in the task). The estimates provided by participants who had been unconsciously primed, however, were uncorrelated with their actual cooperative behaviors, suggesting that they had been pursuing the cooperation goal without knowing they were doing so at the time.

Other research has consistently linked goal priming with downstream effects on judgment and behavior that are characteristic of a person in a motivated state. Phenomenal qualities once considered exclusive to conscious, deliberate goal pursuit, such as persistence in the face of obstacles, increase in goal strength over time, and changes in mood and goal strength depending on the fate of the goal attempt (Atkinson & Birch 1970; Bandura 1986; Gollwitzer & Moskowitz 1996; Goschke & Kuhl 1993; Heckhausen 1991; Lewin 1926) have been demonstrated as well for unconsciously operating goals (e.g., Aarts et al. 2007; Bargh et al. 2012; Bargh et al. 2010; Bongers et al. 2009; Custers & Aarts 2005; Custers et al. 2008; Dijksterhuis & Aarts 2011; Ferguson 2008; Förster et al. 2007).

In their review of goal-priming studies, Custers and Aarts (2010) noted that the traditional determinants of *whether* a conscious goal is pursued, *attainability* and *desirability* (e.g., Austin & Vancouver 1996), also increase the chances that an unconscious goal will be pursued. First, priming the mental representation of a goal activates the skeletomotor impulses connected to execution of that particular action; the researchers note that this essentially "readies" a person for pursuit, thus increasing the attainability of that goal (i.e., facilitating action on time-limited opportunities).

A wide variety of situational features have been shown experimentally to prime or unconsciously activate relevant goals, from social contexts such as having power (Chen et al. 2001; Custers et al. 2008), to material objects such as dollar bills or briefcases (Kay et al. 2004; Vohs et al. 2006), scents (Holland et al. 2005), and even the names of significant others in one's life (Fitzsimons & Bargh 2003; Shah 2003). In the everyday world, the presence of a goal-relevant object usually signals the presence of an opportunity for pursuit (e.g., when a person encounters a piece of cake, usually, he or she has an opportunity to eat it). The context-sensitivity of goal activation highlights how goal processes can unconsciously prepare a person for pursuit the instant that potential opportunities arise.

Second, Custers and Aarts (2010) note the mental representation of a goal can become associated with positive affect; this "tagging" signals goal desirability, which subsequently guides individual behavior toward the desired end-state. In one experiment, participants were exposed to either goal (puzzle-related) words subliminally paired with either positive words (e.g., *friend; beach*), negative words (e.g., *garbage; disease*), or neutral words (Custers & Aarts 2005). Participants were most interested in completing puzzles if they had been seen the puzzle-related words subliminally paired with positive words. Conversely, subliminally pairing goal representations with negative affective stimuli decreased the effort people expended toward associated end-states (Aarts et al. 2007; for similar results with subliminal monetary reward cues, see Schmidt et al. 2010).

Theoretically, a self-perpetuating cycle of goal pursuit could occur under certain circumstances. Similar to how a successful selfish gene replicates itself into future generations (Dawkins 1976), goal representations become more likely to be pursued in the future, through mechanisms that operate independent of individual awareness or guidance. Once a goal representation becomes associatively linked to the situations in which it is frequently and consistently pursued (Bargh 1990), this link increases the likelihood that the goal will be pursued through those specific means instead of others (e.g., Bargh 1990; Gollwitzer 1999; Veltkamp et al. 2008). Similarly, Strack and Deutsch (2004) concluded that need states become strongly linked in memory with the behaviors and situations in which the need was satisfied, so that subsequent experiences of that need activate a bias in attention, or perceptual readiness (Bruner 1957) for those same situational cues. Hence, unconscious goals may become active at precisely those situations in which the individual has historically attained the goal, such that with each successful attempt, the original cues become more reliable triggers of pursuit in the future.

Affective processes may also reinforce the pursuit of successful unconscious goals. Research suggests that when people pursue conscious goals successfully (versus unsuccessfully), naturally occurring positive (versus negative) affect produces a strengthening versus weakening of that goal-pursuit tendency (selecting that goal over other possibilities) in future situations (e.g., Bandura 1977; 1986). A similar positive conditioning effect occurred during the selection of unconscious goals, as shown by Aarts et al. (2007). Consequently, it is possible that a snowballing or virtuous-cycle effect is possible wherein goals activated unconsciously by environmental cues that are successful in their pursuit will naturally become stronger (i.e., more likely to be pursued in the future), whereas those not as successful became weaker based on performance feedback (success vs. failure) from the environments frequented by the person.

4.2.2. Dissociation between active goal operation and individual awareness. Neuroscientific evidence also supports the dissociation of action systems from awareness. That executive control structures can operate without the person's awareness of their operation would require the existence of dissociable component processes within executive control or working memory structures (Baddeley 2003; Baddeley & Hitch 1974; Buchsbaum & D'Esposito 2008). Evidence of such dissociations has been reported in stroke patients with "environmental dependency syndrome" caused by lesions in the frontal cortical lobes (Bogen 1995; Lhermitte 1986). The behavior of these patients was almost entirely driven by situational cues – for example, gardening in a public park (for hours) upon sight of a rake, drinking a glass of water every time the glass was filled despite complaining about being painfully full – with the patients entirely unaware of the oddity and irrationality of their behavior.

Moreover, patients with lesions in the ventral-visual system are unable to consciously report the visual properties of an object, yet are able to incorporate the inexplicable information into movements to grasp the object; whereas those with lesions in the parietal lobe can report on the object but are unable to reach for it successfully (Goodale et al. 1991). Brain damage associated with *anarchic hand syndrome* results in autonomous, goal-directed movements of a limb that cannot be consciously inhibited (e.g., one patient, M.P., reportedly would use one hand to select a channel on the television remote while the other hand – against his will – would immediately press another button; Marchetti & Della Sala 1998).

This evidence has led some to conclude that conscious intentions are represented in the prefrontal and premotor cortex, whereas the parietal cortex houses the representation used to guide action (Frith et al. 2000). (For additional evidence of the operation of action systems dissociated from conscious awareness, see Dijksterhuis & Aarts [2011], Milner and Goodale [1995], Morsella & Bargh [2011], and Wegner [2002].) We take such findings as additional support for the notion that the mechanisms guiding individual behavior evolved separately from the mechanisms furnishing conscious awareness of their operation.

4.2.3. Unconscious goal conflict. Moreover, to the extent that individual behavior was originally driven by multiple, unconscious goals, research should reveal the presence of goal conflict that occurs in the absence of awareness. Kleiman and Hassin (2011) replicated procedures and findings from the earlier goal-priming studies (Bargh et al. 2001) to study whether cooperation and wealth accumulation goals could come into conflict with each other, while the person remains unaware of this tension. Participants were exposed to either cooperation-related or goal-neutral words in an ostensible language test before completing a Commons Dilemma exercise in which participants' usual responses are to compete and thus to accumulate wealth (although they could cooperate instead). The experimenters reasoned that unconsciously priming participants with the goal to cooperate would give rise to goal conflict. Consistent with this prediction, goal-primed participants manifested signs of goal conflict, displaying increased arousal through higher skin conductance levels, requiring more time to reach their decisions, and vacillating more between cooperative and selfish behaviors, as compared to control participants. Moreover, these implicit markers of goal conflict persisted over a five-minute delay and were uncorrelated with participants' self-reported experiences of conflict, thus supporting the notion that conflict can be goal-driven but can occur outside of awareness.

Yet another prediction that follows is that goal conflicts can be resolved independently from conscious guidance. Although social psychological research has yet to directly address this question, there is reason to believe that unconscious conflict resolution is possible. As described above, psychological research has revealed unconscious mechanisms that function to preemptively *minimize* goal conflict. An active, unconscious goal may dominate other goals by superseding conflict altogether, similar to the manner with which people physically remove temptations from their surroundings to bolster their chances of successful self-regulation (e.g., Kuhl 1984; Kuhl & Weiss 1994). Indeed, a phenomenon called *goal shielding* demonstrates how an active goal can unconsciously interfere with how opposing goals come to mind (and thus subsequently influence perceptions and behavior; Shah et al. 2002). In experiments on this topic, participants are slower to recognize words related to one goal (e.g., *diet*) if they are first subliminally primed with a word related to a competing, yet desirable goal (e.g., *cake*, which is eaten for enjoyment; Shah et al. 2002; see also Fishbach & Shah 2006; Veling & van Knippenberg 2006).

Research also suggests that people who are good at self-regulating are particularly effective at inhibiting temptation-related stimuli: these participants are actually faster to recognize goal-related words (e.g., *work*) if those words are subliminally preceded by temptation-related words (e.g., *basketball*; Fishbach et al. 2003). According to Counteractive Control Theory (Fishbach & Trope 2005), temptations preemptively bolster people's higher-order goals and diminish the appeal of the temptation, thereby driving effective self-regulation. Such findings and theories are in harmony with the premise that unconscious goals are associated with mechanisms that negotiate conflict in the absence of awareness. Stated another way, self-control (or control of the self in situations of potential goal conflict) can be exerted unconsciously.

In the study of complicated goal systems, social cognitive experiments may best contribute by documenting phenomena observed in the activity of a single goal (or a few goals), whereas other areas of science (e.g., computer science) may be better suited to examine how goals interact within dynamic, larger systems. We offer here a preliminary conjecture regarding the resolution of goal conflict that, following a period of goal conflict (signified by processes similar to those documented in Kleiman and Hassin [2011]) and resolution, the output might involve phenomena described by Counteractive Control Theory, wherein the strongest or most motivating goal eventually "wins out" in the self-control conflict and eventually steers the individual's behavior within that particular situation or in others like it.

4.3. The reconfiguration principle

The reconfiguration principle predicts that the downstream constellation of goal-priming effects observed in social cognitive research can be understood as constraining the person's cognitive and affective machinery for the purposes of facilitating goal pursuit. A stronger version of this hypothesis holds that the active goal is powerful enough to reconfigure that mental machinery, to the point of making typically effortful processes efficient and automatic if this is necessary for successful goal pursuit, and by inhibiting the normally chronic, automatic effects if they would serve to interfere with successful goal completion.

4.3.1. Attention and perception. Once active, the goal directs one's attention toward some (i.e., goal-instrumental) stimuli and away from others; in effect, the world is seen through goal-colored glasses. Salient, unusual events can be missed entirely when a person is pursuing a goal, as in the "inattentional blindness" research (Mack 2003; Simons & Chabris 1999). In one well-known study, many participants who were busy with the explicit, conscious task of counting the number of ball tosses between actors on a computer display actually failed to notice a gorilla walking right through the ball-tossing game.

Indeed, similar to its conscious counterpart (Anderson & Pichert 1978; Hastie & Park 1986), an unconsciously operating impression formation goal causes greater selective attention to behavioral information inconsistent with the target's general and emerging pattern of behavior (Chartrand & Bargh 1996; McCulloch et al. 2008). Highly accessible goal constructs provide "orienting value," automatically guiding the individual's attention to relevant stimuli in the environment (Bruner 1957; Roskos-Ewoldsen & Fazio 1992), which increases the probability that these objects will be used to achieve that goal.

Goal-facilitating objects can also appear more accessible along different dimensions, for example, by appearing closer in proximity to the pursuer or even larger in size. Balcetis and Dunning (2010) demonstrated that people perceive the spatial orientation of desirable objects (which are the objects that help people achieve their goals; Ferguson & Bargh 2004) as being closer, compared to undesirable objects that are the same actual distance away. Similarly, Veltkamp and colleagues (2008) established that participants who were subliminally primed with a gardening goal overestimated the size (height) of goal-instrumental objects (e.g., a shovel), but not of goal-irrelevant objects (e.g., a pen).

These recent findings are consonant with a long-standing research tradition in social psychology on the motivated perceptual interpretation of events (e.g., Kunda 1990). For example, at sporting events, fans of both teams involved are certain that the referees are against them (Hastorf & Cantril 1954), and people consider studies that produce evidence consistent with their existing beliefs (e.g., concerning the death penalty) as being objective and properly conducted, whereas those studies producing findings contradicting their beliefs are believed to be flawed and biased (Lord et al. 1979). Self-protective motivations transform ego-threatening outcomes (e.g., failure on a test) into more palatable versions that place the blame on external factors (e.g., blaming others; Crocker & Park 2004). One might consider one's teacher to be fair and competent until she gives one a bad grade, and thereafter one might consider her to be biased and incompetent (Sinclair & Kunda 2000). In each of these examples, information processing occurs in the service of the goal, regardless of whether the individual is aware of it.

4.3.2. Evaluation. People's everyday judgments of other people, objects, and events are strongly influenced by how those stimuli relate to the goals they are pursuing. This principle has been a staple of social and motivational psychology since the seminal writings of Kurt Lewin (1935, p. 78) who defined the *valence* of an environmental object or event in terms of whether it helps or hinders the attainment of one's current goals and the satisfaction of one's current needs. Experimental findings continue to substantiate Lewin's dictum (Ferguson & Bargh 2004). In one such experiment involving the sequential priming paradigm (Ferguson 2008), participants were subliminally exposed to either cooperation goal primes (*giving, nice*) or control primes. Afterwards, their implicit positivity toward either goal-relevant (*help* and *generous*) or control words was measured. Participants who were subliminally primed to cooperate displayed increased positive attitudes toward goal-instrumental words but not control words; and in a subsequent experiment, participants' implicit positivity toward goal-related primes predicted the degree to which they actually helped another person.

Active goal influence is so powerful that it can change evaluations of friends, enemies, and even significant others – the very people about whom one's opinions presumably remain stable over time. Fitzsimons and Shah (2008) found that participants who were unconsciously primed with an achievement goal evaluated friends who had helped them with their academic pursuits more positively compared to friends who had not helped them academically. This momentary favoritism toward goal-instrumental friends was not observed for unprimed control participants (see also Fitzsimons & Fishbach 2010).

The goal-driven nature of these interpersonal evaluations is reminiscent of the successful intervention by Sherif and colleagues (1961) in the classic "Robbers' Cave" study. The Rattlers and the Eagles, two warring groups of boys at a summer camp, antagonized each other with increasing violence until they were given an important shared goal. In a situation where everyone's cooperation was needed (e.g., freeing a truck that was stuck in the mud to get food for the entire camp), a Rattler's help became instrumental for an Eagle's goals (and vice versa). Changing the campers' goals dramatically changed how Rattlers and Eagles perceived one another and transformed summer-long rivals into close friends.

4.3.3. Overriding automatic processes that conflict with the active goal pursuit. The transformational power of the active goal over cognitive and affective processes is further indicated by its ability to override otherwise chronic, automatic encoding tendencies. For example, there is much evidence of the automatic manner in which other people are automatically encoded or categorized in terms of their race, age, and gender (e.g., Bargh 1999; Brewer 1988). Recent research, however, suggests that chronic goals assumed to be egalitarian inhibit the same prejudicial biases previously assumed to be automatic and uncontrollable (e.g., Kunda & Spencer 2003; Maddux et al. 2005; Moskowitz et al. 1999).

A similar overriding effect of automatic, prejudicial processes occurs with temporarily active goals as well. Macrae and colleagues (1997) found that giving participants a task goal to detect the presence or absence of dots on facial photographs eliminated any automatic stereotype activation effects when minority faces were shown (since stereotype activation was irrelevant to participants' processing objective). Research also suggests that default negative racial responses to African-American faces on the Implicit Association Test (IAT) can be flipped into positive evaluations when participants are informed that those same faces belong to their online teammates (Van Bavel & Cunningham 2009). These findings are consistent with the notion that joint goals (which are introduced by new alliances) can override automatic processes, causing the recategorization of out-group members into in-group members.

4.3.4. Creation of temporary automatic processes. Goal operation can give rise to novel and temporarily automatic effects as well. For example, implementation intentions, in which one commits oneself to a goal-furthering action in advance by mentally associating a specific concrete goal-pursuit action with an expected future event ("when, where, and how" the action will take place), have been shown to be highly effective means to attain otherwise difficult ends (diet, exercise, difficult health regimens; Gollwitzer 1999; Webb & Sheeran 2006). Implementation intentions effectively delegate control over one's future behavior to the environment, so that a specified, reliably occurring (e.g., routine) future event becomes the automatic trigger of that desired behavior (Gollwitzer 1999). In this way a temporary or strategic automatic effect is created in the service of conscious goal pursuit.

Many so-called automatic effects are in fact goal-dependent (Bargh 1989), in that they only occur in the context of a specific active goal, such as in many automated acquired skills (e.g. Shiffrin & Schneider 1977). Driving a car is commonly understood as an automatic skill for the experienced driver, but this depends on the individual having the goal of driving somewhere in the first place. Stimuli that produce an immediate and unintended reaction under one active goal do not do so under a different active goal, such as when driving a car and kicking one's right foot out to hit the brakes upon suddenly seeing a red light, but not kicking out that foot when seeing the same red light while walking on the sidewalk (Asch 1952, p. 105; Bargh 1992).

Spencer and colleagues (1998) provide perhaps the most dramatic example of a nonautomatic process becoming automatic when it facilitates the current goal pursuit. Research suggests that conditions such as attentional load can prevent people from engaging in negative stereotyping processes. Spencer and colleagues (1998) reasoned, however, that negative stereotyping is a means through which one can enhance one's own self-esteem (at the expense of others), and therefore should persist even in conditions that normally impede stereotyping effects given participants' active needs to restore their self-esteems. Indeed, by providing (bogus) feedback that participants had done very poorly on a task, the experimenters were able to elicit automatic stereotyping effects under conditions where such processes normally do not occur, thereby providing a particularly powerful demonstration of the active goal's ability to "selfishly" reconfigure a person's cognitive machinery in the service of its own pursuit.

4.4. The similarity principle

If conscious and unconscious goal pursuits make use of the same underlying motivational circuits and system, one would expect a high degree of similarity between conscious and unconscious goals in terms of the various component processes recruited, as well as the outcomes produced. An even stronger hypothesis is that conscious goal operation will share the autonomy of operation that is necessarily observed with unconscious goals (because the latter operate outside of conscious self-control). Thus, even when an individual is consciously pursuing a goal, that person is not necessarily in control of or even aware how engaging in goal pursuit has transformed his or her behavior and experience of the world. Recent experiments directly bear on this prediction.

4.4.1. Similarity in processes and neural structures. As previously mentioned, in the original goal-priming studies, unconsciously operating goals produced similar outcomes as when those same goals are pursued consciously, as well as with the same phenomenal qualities (Bargh et al. 2001; Chartrand & Bargh 1996). Since then, research continues to demonstrate similarities between both varieties of pursuit, including the recruitment and use of similar affective processes (Custers & Aarts 2010).

Lending further credence to the notion that conscious and unconscious goals operate using similar processes, McCulloch and colleagues (2008) demonstrated that a primed, unconsciously operating goal to form an impression of a target person follows the same processing stages as was long known for the case of conscious impression formation (see Hamilton et al. 1980; Srull & Wyer 1979; also Chartrand & Bargh, 1996). Compared with a nonprimed control group, in separate experiments, priming the impression formation goal caused participants to (a) be faster to encode behaviors in trait-categorical terms, (b) be more likely to form associations between behaviors, and (c) notice and remember impression-inconsistent behaviors, all well-established subprocesses of conscious impression formation.

Indeed, cognitive neuroscience studies of the brain regions involved in motivated behavior support a model wherein the same underlying mechanisms govern both unconscious and conscious forms of goal pursuit. Pessiglione and colleagues (2007) showed that people automatically increased effort on a hand-grip task when the reward cue (amount of money to be won on that trial) was presented subliminally, the same as what occurred when the reward cue was presented to conscious awareness. They also found that the same region of the basal forebrain moderated task effort level in response to both the consciously perceived and the subliminally presented reward signals. The authors concluded that "the motivational processes involved in boosting behavior are qualitatively similar, whether subjects are conscious or not of the reward at stake" (Pessiglione et al. 2007, p. 906).

Mainstream accounts of executive control or working memory within cognitive science long held that all of the contents of working memory were accessible to conscious awareness – indeed, until recently, "working memory" and "conscious awareness" were used as synonymous terms (e.g., Smith & Jonides 1998). Yet for goal pursuits to operate unconsciously, in real-time interaction with the fluid and dynamic external environment, active goals must make use of flexible working memory structures that operate on and often transform incoming informational input to serve the goal's agenda (Cohen et al. 1990).

Participants in studies in which goals are primed and activated unbeknown to them cannot know in advance which goal-relevant stimuli might be presented; in fact, they are not even aware of which stimuli are goal-relevant and which are not. Nevertheless, in each experimental demonstration of unconscious goal pursuit, the primed goal produced the goal-appropriate outcomes, just as with conscious goal pursuit. For the obtained results to have occurred, the active goal had to be ready for whatever goal-relevant environmental input might arise and then operate on it when it did occur; unconscious goal pursuit therefore must involve the use of executive control and working memory functions as used in conscious goal pursuit (Frith et al. 2000; Hassin 2005).

Direct evidence on this point has been provided recently by Marien et al (2012). In six experiments, the researchers subliminally primed a variety of different goals (e.g., socializing, academic performance) and demonstrated that the unconsciously activated goal "hijacked" the executive control functions of the mind, as revealed by their taking attentional capacity away from an ongoing conscious task such as proofreading.

In this way, unconsciously operating goals can produce the same flexibility of responding to a given set of stimuli that traditionally has been considered the exclusive province of conscious processes (cf. Morewedge & Kahneman 2010).

4.4.2. Autonomy of conscious goal operation. If conscious and unconscious goals are similar, then the autonomy with which unconscious goals operate (which is clearly demonstrated when unconscious goals reconfigure judgment and behavior, with the individual unaware of the goal and thus unable to guide its progress) should also characterize the pursuit of conscious goals. That is, even the goals one intends to pursue, and of which one is aware, are capable of producing information processing and behavioral effects consistent with the goal's agenda but not necessarily with the individual's self-related values and/or overall interests, thus potentially leading to unintended consequences of intended (conscious) pursuits.

Accordingly, Bargh et al. (2008) hypothesized that conscious goal pursuit would operate on any relevant or *applicable* (Higgins 1996) information in the environment regardless of whether the individual intends or is aware of this operation. Participants watched a videotaped interaction with the goal of evaluating the interviewee's suitability for a specified job – either as a restaurant waiter or a newspaper crime reporter. These jobs were included because the desired personality characteristics of a waiter – deferential and polite – and a crime reporter – tough and aggressive – are opposites of each other. In a control condition, participants were told merely that they were watching two people getting reacquainted.

During the taped interview, the two conversation partners were interrupted by a person named Mike, who behaved either politely or rudely. After viewing the tape, participants were given a surprise impression task in which they rated Mike, and not the job candidate on whom they had consciously focused. Not surprisingly, participants in the waiter and control conditions liked the interrupter significantly more if he were polite than if he were rude, but participants in the reporter-goal condition actually showed the reverse preference: They liked rude Mike more than polite Mike. Intuitively, these results may appear surprising. As shown by the ratings of participants in the control condition, people do not normally find rude, aggressive people likable. This research suggests, however, that people will actually like such offensive people if such traits are valued within their currently active goal pursuits – even if the consciously intended focus of that goal pursuit had been another person entirely.

In the next section, we explore how conscious goals not only operate, but also turn off in ways that create inconsistencies in a person's behavior over time. The following evidence for the inconsistency principle is thus further evidence for similarities in operation and outcomes between conscious and unconscious goals.

4.5. The inconsistency principle

As multiple goals within a single individual operate autonomously, becoming activated, operating, and turning off with achievement of their associated end-states, a person's behaviors and judgments will continue to vary as a function of which goals are most motivating. To an outsider, that person's behavior may appear inconsistent over time and, at extremes, even contrary to his or her general self-interests.

As previously mentioned, multiple goals operate by changing a person's perceptions and behaviors in ways that encourage the attainment of their own end-states. A person may pursue a goal even if doing so jeopardizes his or her physical health (Griskevicius et al. 2009; Hill & Durante 2011), and even to the point of feeling helpless to prevent the goal pursuit (Loewenstein 1996). Indeed, many of the ways through which goal operation influences information processing and behavior (i.e., operating outside of a person's awareness and guidance; changing how he or she perceives and behaves in the world) help explain why that person's actions may map inconsistently, weakly, or at times not at all, onto what is clearly beneficial for that person.

4.5.1. Effects of goal turn-off. Attainment of the goal's end-state can also produce inconsistent effects for the individual holding that goal, as when goal completion leads to behaviors that are contrary to expressed important values. In the "goal turn-off effect," once a goal pursuit attempt is completed, the goal deactivates (e.g., Atkinson & Birch 1970; Lewin 1926) and then for a time inhibits the mental representations used to attain that goal (Förster et al. 2005; Marsh et al. 1998), theoretically in order to allow other important goals to be pursued. When a goal is achieved, its downstream influence on cognition and behavior disappears for a time, which can ironically produce behaviors contrary to those originally encouraged by that goal.

For example, research on "moral licensing effects" (Monin & Miller 2001) demonstrates how the operation and completion of conscious goals can produce judgments that can appear inconsistent with that individual's recent behaviors or self-professed values (also Bargh et al. 2008). In one study, participants who were given the opportunity to disagree with blatantly sexist comments were ironically more likely than a control group to recommend a man than a woman for a stereotypically male job (Monin & Miller 2001). Similarly, in another study, supporters of then-U.S. presidential candidate Barack Obama were first given the opportunity to publicly endorse him (or not, in the control condition; Effron et al. 2009). Afterwards, all participants judged the suitability of a job for white versus black people and allocated funds to organizations serving white or black people. Compared to the control group, endorsing Obama caused participants to rate the job as more suitable for whites than blacks and to allocate funding to white causes at the expense of black causes. This latter effect held only for those participants who had scored high on a measure of racial prejudice.

4.5.2. Relation to self-deception phenomena. The capability of active goals to operate and become completed independently from the individual's conscious desires can be seen as "self-deceptive' insofar as the individual remains unaware – or inconsistently aware – of his or her motives. If participants do not know the actual reasons for their behavior when it is influenced by unconscious means (such as through priming manipulations), they should be prone to misattribute the reasons for their behavior to plausible (oftentimes desirable) reasons that are accessible to their conscious awareness. Just such an effect was reliably demonstrated in a series of studies by Bar-Anan et al. (2010). Compared to a control condition, male participants primed with the mating goal were more likely to choose to work with a female tutor on Topic A than a male tutor on Topic B; however, they later explained their choice in terms of greater interest in that topic (which had been randomly paired with either the male or female tutor).

Recent theories identify occurrences of self-deception as an evolutionarily adaptive strategy (e.g., McKay & Dennett 2009; von Hippel & Trivers 2011), because if one is not aware of pursuing a particular goal, one will be less likely to display subtle but telling cues of lying (such as looking nervous or being cognitively overloaded). The Selfish Goal model suggests an alternative, or at least an additional, perspective: To the extent that human judgment and behavior were driven by goal processes before a central "self" even evolved, many instances of "self-deception" can be seen as a result of the autonomous nature of all goal pursuits. Both conscious and unconscious goals encourage single-minded pursuit of the end-state and are capable of producing effects that appear on the surface to be in the service of "deceiving" the individual, such as a "study buddy" being a best friend before, but not following, a big exam.

5. Conclusions

When viewing a person's judgments across time and situations, observers tend to see that person as a coherent whole – as a single agent selecting behaviors expressed in a complementary single body. In this article we argued for a change in perspective. Our central claim here is that an individual comprises multiple goals, each of which exerts a "selfish" influence on how that person sees the world and behaves in it, guiding judgments and behavior in the service of the current goal but not necessarily in the service of the individual's actual, overall best interests.

The Selfish Goal model predicts widespread *automaticity* of higher mental processes and *reconfiguration* of a person's perceptual and behavioral processes according to which goals are most motivating in the current situation. Consistent with the notion that they guided behavior in the evolutionary past prior to consciousness, goals operate by steering an individual's behavior toward goal-specified end-states, even in the absence of that individual's awareness or consent. Hypothesized *similarities* between conscious and unconscious goals led to the observation that, indeed, both forms of pursuit recruit similar processes and are capable of operating and turning off without guidance or even awareness of the individual. As a person temporarily acts in alignment with one goal and then suddenly does not when that goal is completed, *inconsistencies* to behavior as well as unintended effects arise.

5.1. Future directions

Important future directions include clarifying the relationship between the present model and contemporary motivational models in psychology, particularly as they apply to conscious and unconscious goals. Although we have highlighted the similarities between conscious and unconscious goals, recent work is focusing on the key differences between varieties of thought. Notably, however, one recent review has concluded that the difference between conscious and unconscious processes is not in the relative role played in the guidance and production of behavior (Baumeister & Masicampo 2010), consistent with our present argument that much of behavior is governed by unconsciously operating mechanisms. The growing consensus appears to be that conscious processes are superior in serving integrative functions (see Morsella 2005 regarding phenomenal awareness; Baumeister & Masicampo 2010 regarding conscious thought), especially regarding certain types of information such as simulating future scenarios and the perspectives of other people (Baumeister & Masicampo 2010; see also Dijksterhuis & Aarts 2011). This is clearly a fruitful avenue for further comparative research in order to determine even more precisely the uniquely important functions we gain from conscious processing capabilities.

Given the generativeness of the analogy professed herein between genes and goals, another direction for future research involves understanding the precise nature of that link. Indeed, Ernst Mayr (1976) argued that this link was critical for the expression in present time of genetic influences from the very distant past: how genes both pre- and post-natally influence the content, emergence, and structure of programs that guide adaptive attentional, judgmental, motivational, and behavioral responses to one's environment.

5.2. Implications

Today, just as did Freud (1901) more than a century ago in *The Psychopathology of Everyday Life*, contemporary psychological theorists are invoking the concept of motivation (unconscious or conscious) in their explanations for why people behave in ways that seem to run against their self-interest and values. The Selfish Goal model offers potential theoretical insight and implications for motivationally based research at multiple levels of analysis.

For example, people often experience self-control failures in health-related domains (e.g., smoking, drug abuse). In these cases, people are aware that certain actions are detrimental to their physical health, yet they somehow fail to act on that available information. Goal-related interventions might highlight the importance of changing the environment to activate intervention-consistent goals and understanding the dynamic relationship between opposing goals in situations of self-conflict (for example, a dieter is more likely to eat calorie-rich foods after exercising; Fishbach & Dhar 2005). Indeed, models of goal conflict have been used to predict people's success at managing the competing goals of eating for enjoyment and dieting (Stroebe et al. 2008). Strengthening the automatic activation of motoric avoidance goals in the presence of addiction-related stimuli is proving an effective method for reducing intensity of cravings for the addictive substance (Wiers et al. 2010). Using strategically automatic implementation intentions, in which goal pursuits are automatically activated at a future point in time by concretely specified (in advance) future situational conditions, is proving a boon to overcoming difficulties in engaging in healthy behaviors across a variety of domains (Sheeran et al. 2013).

Future research might view self or relationally destructive thoughts and behaviors through the lens of autonomously operating goals. Understanding how a goal may become associated with particular behavioral patterns can inform clinical topics as from self-injury to dysfunctional relationships. Indeed, the pervasive and seemingly paramount needs to protect one's self-esteem (Crocker & Park 2004) and to negotiate self-protection and connectedness goals in one's interpersonal relationships (e.g., Murray et al. 2008) ironically lead people to avoid potential sources of social support. Identifying which goals are most motivating for these self-sabotaging individuals and understanding how goal operation can affect self-concepts and interpersonal judgments may provide new directions for therapeutic interventions.

Another intriguing research possibility involves integrating systems-based approaches to the study of goals. For example, goals may encourage behaviors toward their own end-state that can be seen as contrary to the individual's best interests (e.g., self-injurious behavior); but what of the goals that also promote specific social structures? In political psychology, Jost et al. (2008) have focused on *system justification* effects, in which people perceive the current status quo regarding political power and division of resources as legitimate and fair – even those who are low status and for whom the system actually operates against their self-interests. The researchers explicitly appealed to the operation of an unconscious system-justification motive in order to account for these "relatively puzzling cases of conservatism, right-wing allegiance, and outgroup favoritism among members of low-status groups," which can only be understood if they are "not even aware of the extent to which they are privileging the status quo and resisting change" (Jost et al. 2008, p. 596). Thus, viewing human behavior from the perspective of the selfish goal may deepen understanding of social structures wherein individuals behave in ways that seem clearly contrary to their own best interests.

Just as had Dawkins (1976) with the "Selfish Gene" theory, we should emphasize that The Selfish Goal model does not necessitate selfishness at the level of the individual person. The individual or "self" is composed of many goals – self-interested ones to be sure, but prosocial and morally principled ones as well (e.g., Mansbridge 1990; Miller 1999). Studies of children as young as 18 months old suggest that they can represent the intention of their caretakers and use the knowledge to help themselves (Tomasello et al. 2005); that this is an early emerging human capacity supports the idea that it is the active goal that is selfish, not necessarily the person.

Additionally, more communal goals such as cooperation, egalitarianism, helping, and putting the welfare of others over one's own have been shown to operate entirely automatically and unconsciously (Bargh et al. 2001; Moskowitz et al. 1999; Over & Carpenter 2009), demonstrating a *moral* implication of the Selfish Goal: that selfishly operating goals can produce *unselfish* outcomes at the level of the individual person.

As selfishness, in common parlance, means putting one's own welfare and needs above those of other people (Elster 1990; Jencks 1990), one telling demonstration of a "selfless" (at the level of the individual) selfish goal comes from the Chen and colleagues' (2001) study in which participants unconsciously primed with power were given a choice of experimental tasks to complete, with full knowledge that another participant would have to do the remaining tasks. For participants with a communal relationship orientation (who tend to care more about the welfare of those over which they have power; see Clark & Mills 1993), being primed with power actually activated their communal or altruistic goals, causing them, under the influence of having power, to selflessly shoulder more, not less, of the task burden for the other participant. (For other participants without this communal orientation, power indeed had the effect of producing more selfish, self-interested behavior at the expense of others.) Furthermore, these participants subsequently also reported a greater concern with social approval and expressed fewer racist attitudes.

When examining behavior from the perspective of the individual, people can appear inconsistent by thinking, feeling, and acting in contradictory fashion over time. The Selfish Goal model addresses this puzzle, examining behavior from a goal-pursuit perspective. Specifically, behavior selection at the goal level reconfigures the individual's information processing and behaviors, produces person-level behavior that can be unconscious and autonomous, and results in tension between what is "good" for the individual versus what is "good" for the goal. Together, these outcomes explain how contradictions – such as being capable of both selfish and altruistic judgments and behaviors – can arise in a single individual across time and situations.

Yet, these inconsistencies do pose a problem for individuals in a social world in which trust and predictability of behavior are at a premium and are essential for positive, cooperative relations with one's peers (e.g., Fiske et al. 2007; Tetlock 2002). Thus we note in closing that several recent accounts of the purpose of conscious thought have argued that it evolved (was selected for as an adaptive advantage) in order to manage these same public inconsistencies that are produced by selfish-goal operations (Baumeister & Masicampo 2010; Mercier & Sperber 2011; see also Gazzaniga 1984). The conscious self, in this view, is not so much involved in the guidance of our purposive behavior so much as it is in the business of producing rationalizations and socially acceptable accounts for the actions produced at the goal level. Tetlock (2002) has argued that our accountability to others was so important over evolutionary time that we evolved the "politician" (or "defense attorney") social mindset in order to maintain good relations within our group.

In a zinger often credited to Woody Allen, a character in the classic movie *The Big Chill* (1983) remarks that rationalizations were more important than sex, because he had gone months without sex but hardly a day without a good rationalization. That's no small potatoes: not being able to explain or justify any negative outcomes one was involved in to one's peers could come at the cost of ostracism or worse; being able to give a plausible positive account would thus have strong survival value (e.g., Panksepp, 1998). In assigning this valuable politician role to conscious thought and the conscious self, room is thus made for autonomous goal processes as the proximal determinant of human judgment and behavior.

ACKNOWLEDGMENTS
Preparation of this manuscript was supported in part by Grant R01-MH60767 from the U.S. Public Health Service. We are indebted to the editor for patient guidance of the manuscript through the review process, and to the reviewers for thoughtful feedback that substantially improved the paper. We are grateful to past and present members of the Automaticity in Cognition, Motivation, and Evaluation (ACME) lab at Yale for their constant support. A long-term intellectual debt is owed to the pioneering motivation research and theorizing of Peter Gollwitzer and Gabriele Oettingen; in many ways it constitutes the shoulders on which the Selfish Goal model stands, though the arguments and conclusions herein are our responsibility alone. Special thanks for thoughtful feedback provided on previous versions of this paper go to Joshua Ackerman, Lindsey Beck, Margaret S. Clark, Adam Hahn, Spike Lee, Edward Lemay, Ezequiel Morsella, Anna Kaisa-Newheiser, Andy Poehlman, Alexandra Sedlovskaya, Hyunjin Song, Lawrence Williams, Nicholas Wilson, and Eric Uhlmann.

NOTES

1. To Skinner (1953), goals were inferences made by actors or observers about variables responsible for behavior, and thus not a property of the behavior itself. According to this view, knowing why a person is opening a door (e.g., to help a person whose arms are full of groceries) does not enhance objective descriptions of the person's physical behaviors (e.g., grasping the knob between the thumb and fingers; twisting the wrist and pushing with one's body weight).

2. In contrast, a different contemporary definition carves the unconscious along the dimension of awareness (e.g., Dehaene & Naccache 2001; Loftus & Klinger 1992). Unconscious processes are seen as those that operate in the absence of conscious awareness of even the triggering environmental stimuli. In our view, such definitions may inadvertently lead to impoverished views regarding the actual role of unconscious processes in daily life (as in "the dumb unconscious"; Loftus & Klinger 1992). Unconscious evaluative, behavioral, and motivational processes very likely did not evolve to process solely subliminal-strength stimuli; they evolved and were selected for in terms of how well and adaptively they guided overt behavioral responses to stimuli of variable normal intensity and duration.

Open Peer Commentary

Selfish goals must compete for the common currency of reward[1]

doi:10.1017/S0140525X13001933

George Ainslie[a,b]
[a]*School of Economics, University of Cape Town, Rondebosch 7701, South Africa;* [b]*Department of Veterans Affairs, Coatesville, PA 19320.*
George.Ainslie@va.gov
http://www.picoeconomics.org

Abstract: Selfish Goal Theory is compatible with a behaviorally based theory that recognizes mental processes as behaviors. Both envision choices as made by the competition of purposive processes, which are autonomous in that they are not coordinated by an agentic "self." However, the survival of mental processes – termed "goals" or "interests," respectively – depends on a well-documented active mechanism: reward.

Selfish Goal Theory is a step toward bridging the conspicuous gap between cognitive and behavioral psychology that has existed since the "cognitive revolution" of the 1970s (Baars 1986). A key tenet of the cognitive approach has been to avoid positing a directly selective mechanism for choice, such as reward, reinforcement, or utility – a reaction against the behaviorists' total substitution of such mechanisms for explanatory mental constructs (as in Rachlin 1985). Between avoiding selective mechanisms on one hand and avoiding mental constructs on the other, theorists have not developed the obvious possibility that mental processes compete with one another on the basis of prospective reward. Interestingly, Bargh's earlier writings proposed that much of our "thinking, feeling, and doing … is driven by current features of the environment" (1997, p. 2), leading to the suggestion that he was essentially a behaviorist who recognized internal behaviors (Mischel 1997). Huang's and his current proposal interposes a market-like mediating process between environmental features and behavioral output, based on the competition of goals. This is clearly an advance, but it suffers from ambiguity about the basis on which the goals compete.

Huang & Bargh (H&B) analogize the internal selection of goals to the natural selection of organisms. This is an apt comparison, but whereas genes are selected for differential survival by adaptiveness, the selective process for goals is not specified. H&B are clear (and bold, for cognitivists) in stating that an overarching faculty of choice is not necessary – that the conscious self serves mainly a public relations function (sect. 5.2) – but they are vague about their alternative. H&B have it in mind that goals are "incentivized," and they name B. F. Skinner's "operant" as a synonym of "goal" (sect. 4). They note that "representation of a goal can become associated with positive affect" (sect. 4.2.1, para. 5), but this only "signals" goal desirability, rather than creating it. The most important determinant of selection is said to be "the passive activation or *priming* of higher-order concepts by contextual features" (sect. 2, para. 3; authors' italics), but there is no mention of how, in a conflict, some contextual features get more priming power than others.

All H&B's examples of conflicts involve long-term versus short-term payoffs: behaviors that are "contrary to his or her general self-interests" (sect. 4.5, para. 1) such as "dangerously unhealthy" sexual attractants, putting off work and eating tempting foods (sect. 4, para. 3), and "self-control failures" (sect. 5.2, para. 2). Such actions are said not to "map… onto" the person's own well-being, "even to the point of feeling helpless to prevent the goal pursuit" (sect. 4.5, para. 2). The lack of a clear hypothesis about goal selection leads to trouble in explaining why the mapping fails, that is, why contradictory goal pursuits endure over time rather than resulting in simple preference for one over the other. Given the authors' extensive documentation of unconscious processes in all phases of choice, we might have thought that they were building an explanation for such conflicts using Freudian repression. But no. H&B's evidence is that unconscious processes are capable of the same operations as conscious ones, including such self-control procedures as "activation of motoric avoidance goals" against addictive stimuli, and "strategically automatic implementation intentions" to promote healthy behaviors (sect. 5.2, para. 2), so the dimension of consciousness is unimportant. H&B suggest that there is an underlying coherence to these contradictory behaviors, which may only appear inconsistent "to an outsider" (sect. 4.5, para. 1), but they go no further.

On the other side of the behavioral-cognitive gap, a behaviorism that recognizes internal behaviors has no problem with selfish goals. It calls them interests, the set of behaviors (and thoughts) that are based on a particular source of reward (Ainslie 1992, pp. 88–94). Reward is the process that selects for the mental activity that led to it. Far from being a hypothetical construct, reward performs in controlled experiments with mathematical precision (e.g., Daw & Doya 2006) and can be directly visualized with magnetic resonance imagery in specific coordinated brain centers (e.g., Kable & Glimcher 2007). Like the selection of selfish goals, selection by reward has been compared to natural selection (e.g., Gilbert 1972; Vaughan & Herrnstein 1987). The reward process shapes interests – or goals – as a proxy for the evolutionary adaptiveness that must have led to the selection of its mechanism in turn (Ainslie 1992, pp. 179–84). Although the cognitive revolution rightly faulted classical behaviorism for requiring reward to come from external events, cognitivists might not object to endogenous (self-generated) reward. Expected reward is devalued according to a hyperbolic or hyperboloid function of delay (Berns et al. 2007), which, among other effects, could prevent endogenous reward from short-circuiting the behavioral selection process (Ainslie 2013), and might govern "higher-order goals" (sect. 4.2.3). The finding of hyperbolic discounting is especially relevant to the recurrent dominance of "temporarily active goals" such as unhealthy behaviors (sect. 4.1), because it predicts a disproportionately increased forcefulness of goals when they can be imminently fulfilled. As in Selfish Goal Theory, the property of consciousness is not important.

Hyperbolic discounting also predicts a self-control phenomenon not envisioned here: bargaining among separately motivated interests (or goals) that are expected to be dominant at different times, a variant of repeated prisoner's dilemma (Ainslie 1992; 2012). In such intertemporal bargaining, the perception of a current choice as a test case for similar choices in the future recruits incentive for a long-term goal against a short-term one. Like the interaction of autonomous goals, intertemporal bargaining takes place mostly at an intuitive (or unconscious) level, as demonstrated indirectly by thought experiments (Ainslie 2007), and generates familiar ego functions from the bottom up, "eliminating the need for an agentic 'self'" (sect. 2, para. 3). When such bargaining is taking place, rationalization serves to maintain not only public consistency (sect. 5.2), but also intertemporal trust, avoiding the agent's perception of a defection that would undermine intertemporal cooperation.

Thus, two lines of development, one from behaviorism, the other from social-cognitive psychology, lead to a description of the apparent self as a population of autonomously competing processes. However, the two models operate differently. The main determinant of choice in intertemporal bargaining theory is contingent reward, a process that used to be thought of as strictly external and thus too narrow for human behavior. Selfish Goal Theory seems to view the main determinant of choice as priming, a natural variant of hypnotic suggestion (Bargh 2006, p. 155), which has been unrecognized for the very reason that it is largely unconscious. Certainly the great contribution of priming research has been to show how much choice-making occurs without awareness; but this does not mean that the resulting choices escape the influence of prevailing incentives in any major way. Unhealthy behaviors persist because they are attractive in the short run, and, perhaps tellingly, they respond poorly to hypnotic therapies (Barnes et al. 2010). Conversely, hypnotists cannot impose repugnant behaviors on subjects. To account for disruptive impulses and strong defenses, Selfish Goal Theory needs to take account of how goals bargain with each other in terms of a common currency, best characterized as reward.

ACKNOWLEDGMENT

I thank John Monterosso for comments and suggestions. The cited works under the author's control can be downloaded from www.picoeconomics.org. This material is the result of work supported with resources and the use of facilities at the Department of Veterans Affairs Medical Center, Coatesville, PA. The opinions expressed are not those of the Department of Veterans Affairs or of the U.S. government.

NOTE

1. © George Ainslie 2014. This is a work of the U.S. Government and is not subject to copyright protection in the United States.

Fashioning a selfish self amid selfish goals

doi:10.1017/S0140525X13001945

Roy F. Baumeister and Bo M. Winegard
Department of Psychology, Florida State University, Tallahassee, FL 32304.
baumeister@psy.fsu.edu
http://www.psy.fsu.edu/faculty/baumeister.dp.html
winegard@psy.fsu.edu
http://www.psy.fsu.edu/~baumeisterticelab/winegard.html

Abstract: The selfish goal, at some point in evolution, gave rise to a selfish self. In humans, this selfish self might exert influence over goals, deciding upon which to execute and which to inhibit. This, in fact, may be one of the chief functions of the self.

Huang & Bargh (H&B) propose that viewing goals as "selfish," autonomous agents vying for instantiation in the human (and nonhuman animal) nervous system allows researchers and curious observers to make better sense of goal-seeking behavior. Personality psychologists have long bemoaned the fact that people are not reliably consistent from situation to situation (Mischel 2004). H&B contend that such inconsistencies can be explained by attending to the level of the goal within the person rather than to the level of the person. Competing goals jostle for control of behavior, and so different ones take precedence on different occasions, thereby producing inconsistency.

Our interest is in the self. H&B quote Whitman's line "I contain multitudes," but theories of multiple selfhood have never lasted very long. A single person may contain multiple, conflicting goals, yet the person acts "as a single entity (e.g., both feet walking in the same direction) and as a unified moral agent who takes responsibility for his or her actions. It is rarely acceptable for a person to break a promise with the justification that, "I was a different person three weeks ago."

H&B argue that traditional approaches that posit "an agentic, conscious self at the helm, deliberately forming judgments, making decisions about which courses of action to take, and then guiding one's behavior along those intentional lines" (sect. 2, para. 1) are flawed. We agree that such models are inadequate. In this brief commentary, we wish to explore how to revise it (and preserve some shreds of reality for the conscious, agentic self at the helm) within the context of the Selfish Goal model.

Where do goals come from? Some arise from ancient evolutionary systems (and are generated by unconscious processes), especially perhaps ones close to the requirements of survival and reproduction. Becoming hungry does not begin with a conscious decision. Other goals, however, may require conscious assistance. A chess game strategy is typically reasoned out consciously based on rules that were learned consciously. We strongly suspect that the rules of chess cannot be taught unconsciously. (Machines that can learn chess without consciousness are possible only because conscious beings carefully crafted them.)

Hence, perhaps it is most useful to think of the conscious self as consisting partly of the process by which some goals are favored over others. A selfish goal to, say, take a vacation trip to Zurich might find that its best chance for success is to engage conscious processes to plan and arrange the trip. In consciousness, it may have to compete with other goals, such as to save money. It may have to persuade a romantic partner who is hankering toward Aruba, and such discussions invoke selves of persons. (We also suspect that goals do not have interpersonal conversations independent of their persons!) It may have to arrange payment to airlines, and that requires using credit cards that belong to a particular self.

Indeed, the very concept of a selfish goal problematizes the relationship between self and goal. Can a goal be selfish (or behave selfishly) while still being utterly free of and indifferent to all selves? Selfishness would seem to presuppose at least some rudimentary notion of selfhood. The goal of eating is almost never indifferent as far as who eats: My hunger is not satisfied by your munching. Thus, the goal is not eating per se, but eating by a particular person (that is, a self).

In their conclusion section, H&B say it is a fallacy for people to perceive someone as a coherent whole. Yet the selfishness of goals presupposes a coherent whole in some respects, as long as the goal is not indifferent as to which body achieves fulfillment. H&B propose that many goals are “selfless,” but those goals do not seem truly selfless to us. A selfless goal would not care who eats and who goes hungry, or even whose feet the shoes are put on, or whose career is advanced by some particular success. (We assume H&B really did want their own names on their very nice publication!)

The self may well be partly unconscious. As point of reference, at least, a minimal version of self seems indispensable even for much animal behavior. (An animal too may have the goal of eating, but the animal cares very much whether it or another animal eats.) Thus, the argument that goals precede selfhood in evolution is debatable. Some aspects of selfhood seem implicit in most goals. Only the complex, overgrown human self is a late arrival, though H&B are right to question the extent of its pragmatic efficacy.

The idea of the selfish goal thus does not solve but relocates the problem of the unity and emergence of self. The human self is unique in nature and has many new and complex properties. Goals existed long before such a self emerged. Yet the selfishness of goals was presumably there much earlier in evolution, because living things care deeply about the difference between their own survival and that of a rival. (The same holds for reproduction: people are often quite persnickety about specifically who sleeps with whom.) To be sure, the difference between me and you is a minimal form of selfhood – yet one without which much of natural behavior is incomprehensible. The elaborate, remarkable human self is quite another matter, and we concur with H&B that much goal striving occurs without needing advanced, complex selfhood. To us, however, the challenge is to understand how the elaborate structure of the human self evolved out of that minimal version. We suspect, moreover, that it evolved precisely because it facilitated the pursuit and fulfillment of goals – especially in the context of civilized society and multiple, competing goals and constraints.

Selfish goals serve more fundamental social and biological goals

doi:10.1017/S0140525X13001957

D. Vaughn Becker[a] and Douglas T. Kenrick[b]

[a]*Cognitive Science and Engineering Unit and* [b]*Department of Psychology, Arizona State University, Tempe, AZ 85201.*

vaughn.becker@asu.edu

http://www.public.asu.edu/~loids/

douglas.kenrick@asu.edu

http://douglaskenrick.faculty.asu.edu/

Abstract: Proximate selfish goals reflect the machinations of more fundamental goals such as self-protection and reproduction. Evolutionary life history theory allows us to make predictions about which goals are prioritized over others, which stimuli release which goals, and how the stages of cognitive processing are selectively influenced to better achieve the aims of those goals.

> *Reason is, and ought always to be, the slave of the passions.*
> —David Hume, *A Treatise of Human Nature* (1739)

The idea that goals can function as autonomous agents, and that conscious and controlled processes are often merely along for the ride, is a thought-provoking one, with some honored historical antecedents. In his 1739 *A Treatise of Human Nature,* David Hume penned his famous line about reason being the slave of the passions, and in 1986 Marvin Minsky argued along similar lines in his *The Society of Mind*. Critically, though, Minsky also postulated hierarchical organization of control. Goals are not just arbitrarily activated by situational ephemera willy-nilly, but are more likely to be activated according to an evolved system of goal priorities. Being primed with words connoting “elderly” might make us walk down the hall a little slower, but if we suddenly had to run from a hatchet wielding madman, we suspect that the running would be just as fast. Avoiding murder is a mandatory goal, whereas other goals, such as egalitarianism or the avoidance of age-discrimination, are optional luxuries, emerging only when more basic goals are satisfied.

We have argued, based on considerations of evolutionary life history, that goals can be organized into a natural taxonomy (e.g., Kenrick et al. 2010). Fundamental biological and social goals dominate many more ephemeral or easily delayed goals, which can only be selfish when the big players are quiet. Goals like self-protection, social affiliation, and mating are prioritized and can seize control of our behavior even when we are consciously trying to pursue a different task. These goals reflect fundamental motivational systems, which can be conceptualized in terms of Martindale’s (1980) notion of “subselves” – sets of subprograms for dealing with general categories of adaptive problems. This view entails a critical role for environmental inputs. Certain stimuli elicit stronger reactions than others, because they have more significant and/or consistent consequences in the ancestral (or developmental) past. Cognitive systems have thus evolved (or are biologically prepared to learn) a vigilance for stimuli relevant to fundamental goals. Neither the stimuli nor the goals exist in isolation; the psychological system has coevolved with features of the ecology.

By thinking in terms of a hierarchy of evolved, stimulus-specific goals, predictions can begin to address: (1) how goal priorities shift in real-world tasks (which vary systematically with developmental stage and/or ecological threats and opportunities), and (2) how these goals reconfigure cognitive processing to achieve their aims.

Haung & Bargh (H&B) are quite right to emphasize the automaticity of such goals, and their ability to override more controlled processes, but evolutionary theory suggests a pecking order. I might enter the room looking for a colleague, but the presence of a highly attractive person or someone who appears unstable and dangerous will reliably divert my attention from this goal. Of course, we should not expect that a single hierarchy exists for all people at all times. Even within a person, goals shift with time – standing up to an aggressor to impress a potential mate might characterize a man’s modus operandi when young and single, but not at all a few years later when he has a family. On this view, different motivational systems have a natural developmental hierarchy, as depicted in Figure 1. Following Maslow’s scheme, motives lower in the pyramid unfold earlier in life and also take priority over higher goals later in life.

It is the fundamental social goals that will be more likely to reconfigure cognitive processing, doing so in ways that are functionally tuned to the outcomes they seek. Although people typically remember others of the same race better than racial out-group members, for example, that pattern is reversed when people’s self-protective goals have been activated (Becker et al. 2010). Furthermore, although activating self-protective goals leads to better memory for out-group men, it does so without an increase in overt visual attention to those men, because people pursuing a self-protective goal need to encode potential threats without staring at those threatening individuals, lest they invite the very peril they strive to avoid. On the other hand, proximate disease threats lead to the opposite disjunction of attention and memory: we find it difficult to look away from people whose faces appear to have disease symptoms, but this does not translate to greater memory for those people (Ackerman et al. 2009).

In both of these studies, fundamental goals appeared to channel more cognitive effort into one process versus another, but goals

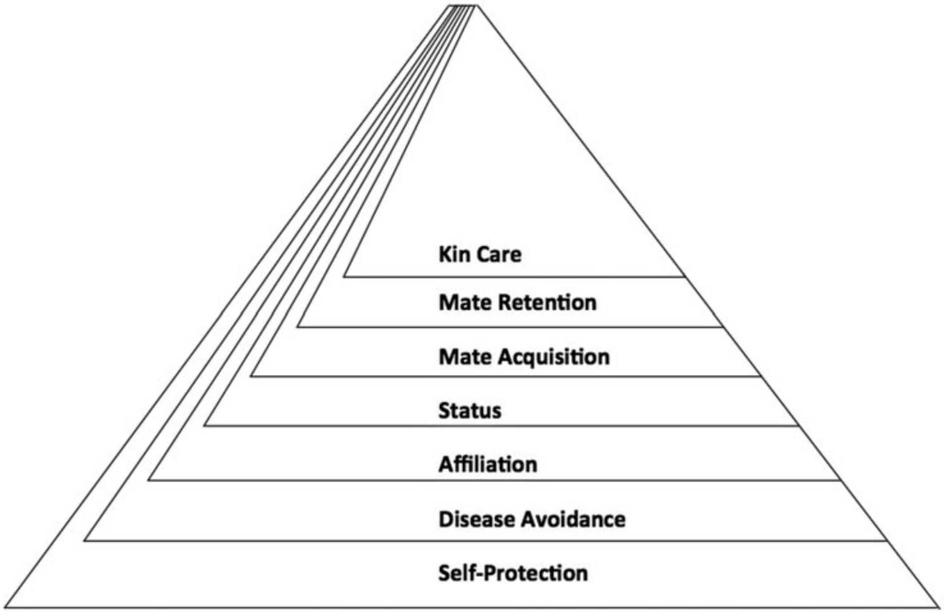

Figure 1 (Becker & Kenrick). A hierarchy of fundamental goals. Following Maslow's classic scheme, those lower in the pyramid are presumed to develop earlier in life, and, at a proximate level, to take priority over those higher up. Following evolutionary life history considerations, the lower goals are linked to somatic effort, followed by mating effort, and finally parenting effort (see Kenrick et al. 2010).

might also more generally boost or undermine accuracy by liberating more cognitive resources. For example, we found that activating the goal of self-protection with a guided visualization enhanced people's accuracy at detecting enemies (Becker et al. 2011). In contrast, activating the goal of revenge/anger (absent self-protection goals) undermined accuracy in favor of a bias to identify slightly angry or foreign-looking faces as enemies. This suggests that fear liberates more cognitive resources than anger, sensitizing perception to maximize the benefits of detecting threats while minimizing the costs of false alarms when no threat is present. One of the most basic findings in social cognition is that people are miserly with their cognitive resources, and the studies above suggest that people may instinctively be saving these resources for occasions when more fundamental goals arise.

In conclusion, selfish goals are a good start, but it is critical to consider how different goals are organized into an adaptive system. The society of mind is just that – a society – it is not a Hobbesian state of nature with all goals being brutish and short with one another. We instead propose a hereditary oligarchy, a hierarchy of fundamental goals that work together to maximize reproductive fitness across highly variable environments.

Tag, you're it: Affect tagging promotes goal formation and selection

doi:10.1017/S0140525X13001969

Eliza Bliss-Moreau[a] and Lisa A. Williams[b]
[a]*Brain, Mind, and Behavior Unit, California National Primate Research Center, University of California, Davis, Davis, CA 95616;* [b]*School of Psychology, University of New South Wales, Sydney NSW 2052, Australia.*
eblissmoreau@ucdavis.edu
http://www.elizablissmoreau.com
lwilliams@unsw.edu.au
http://www.williamssocialemotionslab.com

Abstract: Building upon Huang & Bargh's (H&B's) theory, we propose a complementary view that goal formation and selection are both supported by affect. We suggest that goals may form when affect "tags" discrete behaviors and their outcomes. Further, we propose that goal-associated affect may help guide selection between competing goals, for example, in the case of short-term and long-term goals.

Huang & Bargh (H&B) provide a compelling view regarding the "selfish" nature of goals. Two important questions are left unaddressed. First, how are behaviors prioritized such that they become goals in the first place? Second, how is competition between short-term and long-term goals reconciled? We propose that affect supports both processes. Affect is an omnipresent neuropsychological state generated via the integration of sensory information with interoceptive information; it is characterized by hedonicity and arousal (e.g., Barrett & Bliss-Moreau 2009; Barrett & Russell 1999; Russell 2003). In this commentary, we first propose that affect "tags" behaviors such that they become organized into goals (i.e., goal formation). Second, we propose that affect "tags" existing goals, allowing for the resolution of goal competition (i.e., goal selection).

Affect is broadly involved in many psychological phenomena, including but not exclusive to, emotion (e.g., Barrett 2006; Russell 2003), personality (e.g., Revelle 1995; Yik & Russell 2001), stereotyping (e.g., Mackie & Hamilton 1993; Moreno & Bodenhausen 2001), and the formation of implicit associations (e.g., Payne et al. 2005; 2007). Affect guides behavior (e.g., Martin et al. 1993; Raghunathan & Pham 1999), directs deployment of cognitive resources (e.g., Gable & Harmon-Jones 2010; Wegener et al. 1995), and even influences the content of consciousness (e.g., Anderson et al. 2011b; 2011a). As such, affect is a prime candidate for being involved in goal formation and selection.

H&B acknowledge the possible role for affect in goal processes when they state, "a goal can become associated with positive affect; this 'tagging' signals goal desirability" (sect. 4.2.1, para. 5). In this and similar views (e.g., Aarts 2007), affect perpetuates goals, which presumably already exist. A complementary perspective addresses goal formation. In our view, affect becomes associated with particular behaviors, thus influencing the likelihood that those behaviors occur on future occasions. Goals form over repeated behavior-outcome sequences. For example, the repeated association between physical affection and its associated positive affective state leads to a goal to pursue social relationships that may include physical affection. Thus, affect plays a critical role in goal formation: when positive affect accompanies a behavior-outcome link, that link is prioritized for formation into a goal. Likewise, when a behavior results in negative affect, the likelihood of that behavior occurring in the future decreases (possibly subserving "avoidance" goal formation; Carver & Scheier 1990; Higgins 1997). The linking of affect with behaviors likely occurs via associative learning (Bliss-Moreau & Barrett 2009). These processes can occur either consciously or unconsciously and may become automatized according to the rules set forth by H&B. Our view does not require that goals become mentally represented in a formal sense, though this is likely the case in humans. In sum, we propose that behavior-outcome patterns become organized into goals *because* of affect.

Affect may also help prioritize goals when selection between different goals is necessary (in the "predecisional phase" of goal sequencing; Heckhausen & Gollwitzer 1987). Often, goals are pursued serially rather than in parallel, requiring that a particular goal be selected over another. Goal selection is particularly important when selecting between short-term goals (discussed in the target article) and long-term goals (not discussed in the target article). Pursuing the most desirable goal now (e.g., eating a delicious but not nutritious snack) may negatively impact future goals and states (e.g., losing weight). Thus, competition arises: Should one pursue short- or long-term goals in the given moment? Affect may aid in the resolution of this conflict by "tagging" goals for selection. For example, if a highly positive affective state is associated with a long-term goal, that long-term goal is more likely to be selected than a short-term goal tagged with less positive affect. Affect may also support favoring long-term gain (positive experiences conferred in the future; e.g., losing weight) in the face of short-term pain (negative experiences in the present; e.g., achy muscles after a hard workout) (Williams & DeSteno 2014). Over time, this process likely becomes anticipatory in nature: goals that are anticipated to create more positive states are selected (Bagozzi & Pieters 1998).

We are not alone in suggesting that affect and goals are related. Others have proposed that affect arises from and shapes goal-relevant processes (e.g., Aarts 2007; Carver & Scheier 1990; Schwarz & Bohner 1996). Notably, our focus on the importance of affect's "tagging" role in goal formation and selection guides a number of testable hypotheses. For example, when our perspective is applied to understanding evolution, one could deduce that animals that have affect should also have goals. In this view, it is possible for organisms to have goals without sophisticated cognitive resources or the sorts of mental representations that humans have. Further, the differentiation of affect should track the differentiation of goals. Animals with affect alone (e.g., who feel pleasure) would have relatively more global goals (e.g., moving toward that which produces pleasure) than animals with highly differentiated affect (e.g., discrete emotions such as pride), who should have correspondingly differentiated goals (e.g., achieving personal success). In our view, the process of differentiation occurred across evolutionary history. Analysis of the affect-goal relationship across phylogeny should, therefore, reveal a progressive change from nondifferentiated affect supporting nondifferentiated goals to differentiated affect supporting differentiated goals.

Although cross-species evaluation of this process needs to be performed, it is notable that many theories of discrete emotion in humans utilize similar logic (e.g., Bagozzi & Pieters 1998; Cosmides & Tooby 2000; Frijda 1986; Keltner & Gross 1999). Another testable hypothesis highlights the relatively flexible nature of goals, which is inherently the case as the result of the fluidity of the affect system. For example, it should be possible to shift goal formation and selection by manipulating an individual's affective state. Indeed, support for this idea has been found in humans (Aarts 2007; Schwarz & Bohner 1996).

We hope that focusing on the role of affect in goal formation and selection should allow for greater understanding of the mechanisms that subserve goals in humans and nonhuman animals.

Unconsciously competing goals can collaborate or compromise as well as win or lose

doi:10.1017/S0140525X13001970

Peter Carruthers
Department of Philosophy, University of Maryland, College Park, MD 20742.
pcarruth@umd.edu
http://faculty.philosophy.umd.edu/pcarruthers/

Abstract: This commentary offers a friendly extension of Huang & Bargh's (H&B's) account. Not only do active goals sometimes operate unconsciously to dominate or preempt others, but simultaneously active goals can also collaborate or compromise in shaping behavior. Because neither goal wins complete control of behavior, the result may be that each is only partly satisfied.

Huang & Bargh (H&B) provide powerful arguments for the claim that goals can be active and influence behavior outside of awareness. I concur fully. Indeed, I go further and claim that goals (as opposed to felt desires) are *never* conscious and *always* operate outside of awareness. For our access to our own goals and decisions is always interpretive in nature, much like our access to the goals of other people (Carruthers 2011). H&B focus mostly on the ways in which unconscious goals can *preempt* conscious ones, however, and in their official statement of their view, they propose that behavior is always controlled by the strongest currently active goal. I suggest a complimentary perspective: goals can also cooperate or compromise outside of awareness, simultaneously shaping a single behavior (in cases of compromise in such a way as to partially satisfy each while fully satisfying none). For example, in answering a question, one might have the goal of saying honestly what one believes, but one might also have the goal of making a good impression on the questioner, or of enhancing one's own self-image. As a result, what, precisely, one says may be different from what one would have said had either one of these goals been active individually. I will illustrate and substantiate this claim with reference to the counterattitudinal essay paradigm used extensively by social psychologists studying so-called "cognitive dissonance."

The basic finding in this literature is that participants induced to write an essay arguing for the opposite of what they believe will thereafter shift their expressed attitudes quite markedly (provided that their freedom of choice in writing the essay had been emphasized to them). For example, students who are known to be strongly opposed to a rise in tuition (as measured in an unrelated questionnaire taken some weeks previously, perhaps) will say that they are neutral on the issue, or even moderately in favor, after writing an essay under conditions of "free choice" arguing that tuition *should* be raised. For many years it was believed that writing a counterattitudinal essay induced a negative feeling (called "dissonance") resulting from the perceived inconsistency between one's underlying attitude and one's freely undertaken behavior (Bem 1967; Festinger 1957). But there is good reason to think that this explanation is incorrect. Although the negative emotional component of the account is well established (Elliot & Devine 1994), it turns out that similar shifts in expressed attitude can be caused by *pro*-attitude essay writing, provided that people believe their action is likely to be harmful. This was elegantly demonstrated by Scher and Cooper (1989) who told participants of a newly discovered (but fictitious) "boomerang effect," according to which essays read early in a sequence of messages would tend to have counterpersuasive effects. Hence, an essay arguing *against* a rise in tuition would be apt to induce the university committee dealing with the issue to *raise* tuition if that essay were read first or second in the series of essays they consult when considering the question. Under these conditions people who had written proattitudinal essays (arguing that tuition should *not* be raised) shifted their expressed attitudes quite markedly having learned that their essay would be read second, just as did those who wrote counterattitudinal essays (arguing that tuition *should* be raised) who learned that their essay would be read second-to-last.

The best explanation of these and many similar findings is as follows (Carruthers 2011). People who have had their freedom of choice made salient to them appraise their essay-writing action as having been bad, and this makes them feel bad. When asked later about their attitudes on the topic, they rehearse the behavioral alternatives open to them in the manner of Damasio (1994) and select the one that they appraise as presenting their action as *not bad*, thereby ameliorating their negative affective state. This will often involve saying something other than they believe. Indeed, people will embrace any one of a number of behavioral strategies to rid themselves of negative affect in these experiments, including not only shifting their expressed attitude on the subject matter of the essay, but also denying responsibility for the action or denigrating the importance of the issue. Moreover, they adopt the first such opportunity that is offered to them, and thereafter their responses to the remaining questions are unchanged (Gosling et al. 2006; Simon et al. 1995). As a result, it is implausible that any of their attitudes had really changed in advance of the questions being asked.

What happens, then, when free-choice participants in standard (nonboomerang) counterattitudinal essay-writing experiments are later questioned about their attitudes is this. The question activates their standing attitude (e.g., that raising tuition would be bad) while also activating the goal of saying what one believes, or saying what is true. This goal on its own would lead them to say "Strongly opposed." But they also have the goal of making themselves feel better (or perhaps: the goal of presenting their previous action as having been a good one). This second goal on its own would lead them to say "Strongly in favor" (because in that case their action of arguing in support of a rise in tuition would be appraised positively, and not merely neutrally). But in fact participants tend to answer around the midpoint, thereby partly satisfying each goal while fully satisfying neither. Moreover, it is quite unlikely that either of these goals operates consciously. (For example, participants surely could not be aware of their attitude that raising tuition would be bad, or they would then be aware that their answer is a dishonest one, and this would make them feel worse, not better.)

I conclude, then, that not only do goals initiate behavior in ways that are unconscious, with some goals preempting the activity of others (as H&B claim), but sometimes competing goals can *cooperate* or *compromise* unconsciously; in the latter case, to the partial satisfaction of each.

A deeper integration of Selfish Goal Theory and modern evolutionary psychology

doi:10.1017/S0140525X13001982

Daniel Conroy-Beam and David M. Buss
Department of Psychology, University of Texas, Austin, Austin, TX 78712.
dconroybeam@utexas.edu **dbuss@austin.utexas.edu**

Abstract: Conceptually integrating Selfish Goal Theory with modern evolutionary psychology amplifies theoretical power. Inconsistency, a key principle of Selfish Goal Theory, illustrates this insight. Conflicting goals of seeking sexual variety and successful mate retention furnish one example. Siblings have evolved goals to cooperate and compete, a second example. Integrating Selfish Goal Theory with evolutionary theory can explain much inconsistent goal-directed behavior.

Huang & Bargh (H&B) present a novel meta-theory of human behavior that draws from the success of the genes-eye perspective, the dominant paradigm within modern evolutionary theory. It is inspiring that mainstream psychologists are increasingly acknowledging some of the central tenets of evolutionary psychology. These include: (1) that evolution by selection is the fundamental creative force behind the origins of human psychological mechanisms (Buss 1995; Tooby & Cosmides 1992); (2) that theories of human psychology inconsistent with known principles of evolutionary biology stand little chance of being scientifically correct (Symons 1992); and (3) that because many adaptations, including evolved goals, are somewhat specialized for different functional behavioral output, their manifestations will sometimes be in conflict with each other and individuals will consequently be, or appear to be, inconsistent (Buss 2012; Kurzban 2012).

We propose that the utility of Selfish Goal Theory will be strengthened by even fuller conceptual integration with the principles of evolutionary psychology. Evolution by selection is an essential and logically necessary explanation of the origin of the psychological mechanisms that underlie human behavior. Evolutionary theory provides not merely a metaphor for explaining behavior, but rather an indispensable set of causal principles for explaining why humans have the goals toward which they strive. When properly applied, the genes-eye perspective can be useful in predicting not only specific human goals, but also the "design features" of the underlying mechanisms, including the many properties of goals that H&B describe.

A concrete example from evolutionary psychology illustrates this important point. H&B highlight *inconsistency* in behavior over time as one of the key principles of Selfish Goal Theory. Inconsistency serves as a useful test case for demonstrating the utility of an increased emphasis on evolutionary principles because identifying inconsistencies follows from identifying *specific* goals and their manifestations. A more complete grounding of Selfish Goal Theory in evolutionary psychological principles would facilitate the identification of inconsistency because an evolutionary perspective guides researchers to specific evolved goals, as well as the behavioral inconsistencies that may exist when these goals conflict.

Consider two plausible evolved goals within the mating domain for which there is abundant empirical evidence: (1) the desire for sexual variety (e.g., Schmitt et al. 2003; Symons 1979), and (2) the goal of keeping a long-term mate sexually faithful (e.g., Buss et al. 1992; Daly et al. 1982). Acting on a desire for sexual variety by having an extra-pair copulation seems inconsistent with endorsing moral and political condemnations of adultery and promiscuity in others, which is hypothesized to function in promoting long-term sexual fidelity in one's partner (Kurzban et al. 2010). But these apparent *behavioral* inconsistencies are not *psychologically* inconsistent because they derive from two separate evolved psychological adaptations.

A qualitatively different form of inconsistency highlighted by evolutionary psychology occurs in human sibling relationships. Human siblings share, on average, 50% of their genes by descent, poising sibling relationships to be highly cooperative according to inclusive fitness theory (Hamilton 1964). However, given their 50% lack of genetic relatedness, their similar age, and their shared environments, siblings are also sometimes in competition for major resources such as parental investment, social status, and available mates. These facts combine to suggest that sibling relationships will simultaneously be among the most cooperative *and* conflictual human relationships (Buss 2012). One sibling might rush to the other's aid in a battle with a common enemy at one time, while attempting to monopolize a larger share of parental resources at the expense of the other at another time. The often conflicting evolved goals of investing in close kin and securing resources from shared environments furnish precise predictions about the forms seemingly inconsistent behavior will take.

Conflicting mating goals and conflicting goals within kinship relationships are just two of the many domains in which evolved psychological mechanisms give rise to inconsistency or apparent inconsistency (Buss 2012; see also Kurzban 2012). The key point is that knowledge of evolved goals and their potentially contradictory manifestations is enhanced by analysis of the adaptive functions of goals. A closer conceptual integration of Selfish Goal Theory with evolutionary psychology furnishes the theoretical power required to generate very specific predictions about the domains in which different goals generate inconsistent, or seemingly inconsistent, behavior.

In sum, we believe that Selfish Goal Theory, which draws from modern evolutionary biology and psychology, is an important conceptual step in the right direction. We suggest that a deeper conceptual integration with evolutionary psychology will provide an even richer set of empirical predictions about the ways in which selection has forged the psychological mechanisms that make humans behave in ways that seem highly goal-driven, and the design features of goals that lead to apparent or real behavioral inconsistencies. We hope that other psychologists will follow the lead of H&B and build upon the important first steps their theory provides in creating psychological theories not just consistent with, but explicitly driven by, known principles of evolutionary theory.

Unconscious habit systems in compulsive and impulsive disorders

doi:10.1017/S0140525X13001994

Natalie L. Cuzen,[a,b] Naomi A. Fineberg,[c,d] and Dan J. Stein[a]
[a]*Department of Psychiatry and Mental Health, University of Cape Town, Cape Town 7925, South Africa;* [b]*ACSENT Laboratory, Department of Psychology, University of Cape Town, Cape Town 7701, South Africa;* [c]*National Obsessive Compulsive Disorders Specialist Service, Hertfordshire Partnership NHS University Foundation Trust, Queen Elizabeth II Hospital, Welwyn Garden City AL7 4HQ, United Kingdom;* [d]*University of Hertfordshire Postgraduate Medical School, Hatfield, Hertfordshire AL10 9AB, United Kingdom.*
natalie.cuzen@uct.ac.za **naomi.fineberg@hpft.nhs.uk**
dan.stein@uct.ac.za
http://www.health.uct.ac.za/departments/psychiatry/about/
http://www.humanities.uct.ac.za/hum/departments/psychology/about
http://www.hpft.nhs.uk/

Abstract: It may be useful to consider the application of Huang & Bargh's (H&B's) theory of unconscious motivational processes to psychopathology. In disorders of compulsivity and impulsivity, an unconscious habit system may play a key role in explaining ego-dystonic or self-destructive behaviour. H&B's theory may provide some insights into understanding conditions such as obsessive-compulsive disorder (OCD) and drug addiction; however, additional work is needed to address the neurocircuitry and neurochemistry mediating their abnormal underlying motivational processes.

Huang & Bargh (H&B) propose an interesting theory of unconscious motivation, based in part on evolutionary principles. Early clinical notions of unconscious processes highlighted their importance in compulsive and impulsive disorders, and the target article does refer on occasion to such conditions (e.g., drug addiction). It may, however, be useful to pay additional emphasis to such psychopathology, with particular reference to the question of whether H&B's arguments are valuable in explaining not only normal, but also abnormal motivational processes.

The inconsistency principle, for example, seems particularly apparent in individuals with compulsive and impulsive disorders, such as obsessive-compulsive disorder (OCD) and addiction. Early on Freud emphasized the important contrast between obsessional and hysterical neurosis. More recent work continues to describe how individuals with OCD engage in repetitive behaviour despite acknowledging the absurdity of such compulsions (Foa et al. 1995); similarly, individuals with addiction are unable to cease drug consumption despite being aware that this is not in their bests interests (Loewenstein 1996).

There is growing evidence that an unconscious habit system may play a crucial role in explaining the inconsistency seen in both OCD and drug addiction. OCD may be characterized by an underlying vulnerability to habit formation (Gillan et al. 2013). Although drug addiction may be precipitated by an underlying vulnerability to impulsivity, the addiction process is stimulus-driven and characterized by repetitive, inflexible, and persistent behaviour despite associated negative consequences (Everitt & Robbins 2005; Volkow & Fowler 2000). Put differently, habits restrict agency by prompting automatic responses to environmental cues (e.g., a dirty toilet in OCD, an inviting pub in alcoholism), regardless of whether or not the outcome is detrimental to the individual (Dickinson 1985). Crucially, habits may be a driving mechanism in both avoidant (e.g., OCD) and appetitive (e.g., drug addiction) motivational processes (Gillan et al. 2013). Typically, behavioural control is maintained through a balance between the goal-directed system and the habitual system (Dickinson 1985; de Wit & Dickinson 2009). The appropriation of control away from the goal-directed system toward the habitual system in OCD and drug addiction may be underpinned by anomalies in the frontostriatal circuits governing these functions (Balleine & O'Doherty 2010; de Wit et al. 2012).

Although habitual behaviour in drug addiction and OCD seems to be defined by a relative lack of goal-directedness, H&B's theory of the "selfish" nature of goals may well have some application to abnormal motivational processes in these disorders. Specifically, one may consider habits characterizing OCD and drug addiction to be "selfish," insofar as they involve adaptive systems. Thus, several authors have emphasized that precautionary behaviours and reward-seeking behaviours have an evolutionary basis (Nesse & Berridge 1997).

We suggest that H&B's theory provides some useful insights into understanding compulsive and impulsive disorders insofar as it emphasizes continuity between unconscious and conscious motivational processes, in addition to the notion of the "selfish" or adaptive nature of habitual processes. However, we would argue that additional work is needed in order to address the neurocircuitry and neurochemistry that characterize the relevant motivational processes; there is a good deal of relevant literature (e.g., Everitt & Robbins 2005; Graybiel & Rauch 2000) that may contribute to delineating the precise way in which such processes operate, and which may be of specific value in the treatment of psychopathology.

What's in a goal? The role of motivational relevance in cognition and action

doi:10.1017/S0140525X13002008

Baruch Eitam[a] and E. Tory Higgins[b]
[a]*Department of Psychology, University of Haifa, Mount Carmel, Haifa 31905, Israel;* [b]*Department of Psychology, Columbia University, New York, NY 10027.*
beitam@psy.haifa.ac.il
http://hevra.haifa.ac.il/~psy/index.php/en/faculty?id=88
tory@psych.columbia.edu
http://www.columbia.edu/cu/psychology/indiv_pages/higgins.html

Abstract: We argue that it is possible to go beyond the "selfish goal" metaphor and make an even stronger case for the role of unconscious motivation in cognition and action. Through the relevance of a representation (ROAR) framework, we describe how not only value motivation, which relates to "selfish goals," but also truth motivation and control motivation impact cognition and action.

Huang & Bargh (H&B) present an impressive review of research on unconscious sources of cognition and action. From our perspective, however, in their resolve to clear the path for the "selfish goal" metaphor, they may have missed an opportunity to make an *even stronger* case for the role that unconscious motivational processes play in cognition and action. Here, we outline such a framework (Eitam & Higgins 2010; Eitam et al. 2013) and discuss its implications for the selfish goal metaphor.

The framework, relevance of a representation (ROAR), adds a second, motivational layer – which we called relevance – to one of the basic pieces of conceptual machinery in cognitive and social cognitive psychology: mental representations and their dynamics. The principal tenet of ROAR is that the activation of mental representations (externally or internally stimulated), and hence the accessibility of the concepts they represent, will be a function of the motivational relevance of those concepts.

In line with Higgins' (2012) analysis of motivation, within ROAR we (Eitam & Higgins 2010) named this "goal" source of motivational relevance, that is, obtaining desired results, "value relevance." As H&B convincingly demonstrate, this source of relevance is indeed pervasive. ROAR nicely accommodates the effects of goal activation described by H&B by specifying, for example, how so-called *positive tags* of an action representation (Custers & Aarts 2005) could lead to action selection. ROAR goes further to explain how, although many goal relevant objects exist in the environment, only a few (relevant) goals are activated (see Ferguson 2007) because goals (and their activation) are themselves subject to the same relevance computation as any other knowledge representation.

But pursuing valued end-states, that is, goals, does not exhaust what motivates humans or other animals. Following Higgins (2012), we specified two additional sources of motivational relevance – truth relevance (knowing what is real, what is correct) and control relevance (managing what happens), each of which affects activation independent of value relevance.

Truth relevance implements the mind's need to know what is *really* "out there." If the mind is expecting something (see Bruner's 1957 "perceptual readiness"), once a sufficiently similar signal appears, the corresponding representation will be more strongly activated because of its high truth relevance. Through its effects on mental activation, truth relevance also impacts memory. This has been helpful in explaining the effects of "shared reality" (e.g., Echterhoff et al. 2005). In these studies, participants are asked to read evaluatively ambiguous information about a target person and then present that person to an audience that likes or dislikes that person. Multiple studies have established that, although asked to reproduce the original (ambiguous) description, recall is biased towards the "shared reality" audience-tuned message. When participants are told later that the audience could not (*vs.* could) identify the target on the basis of the description they supplied, this "saying is believing" memory effect disappears. ROAR accounts for this disappearance in the failed communication condition in terms of reducing the truth relevance, and thus the accessibility, of the previously shared, and now "*un*-shared," message.

The third source of relevance we proposed was relevance from control. This source of relevance implements the mind's necessity to know what can be successfully done/affected/controlled by the organism (or by other agents) in the environment. It enables integrating dissimilar results. For example, dopaminergic cells' phasic response is thought to be influenced by the degree of control the stimulus is associated with (Redgrave & Gurney 2006); task (i.e., goal) irrelevant stimuli that are contingent on one's action are nonetheless registered in humans (Band et al. 2009); and action contingent effects are registered in babies prior to any association with goal-directed control (Verschoor et al. 2013).

One could argue that the concept of a goal, selfish or not, could assimilate both the truth and control sources of relevance as being merely highly abstract goals. But this would make the notion of goal representation rather useless as an explanatory construct because it would always be possible to impute an even more abstract goal (such as "be part of God's universe"). Moreover, for any valued goal, there are still motivational questions regarding how real or imaginary that goal is (truth) or how its pursuit can be managed (control), and thus the need for distinguishing value relevance from truth and control relevance remains (see Higgins 2012; Silvestrini & Gendolla 2013).

We want to stress that in ROAR relevance is not a metaphor. Indeed, there are indications that a number of brain structures are involved in computing such relevance (for a review, see Eitam & Higgins 2010). Importantly, the computation of relevance is affected by the organism's current state, including current needs and current knowledge. Thus, even if an activation goal may be selfish once activated, *whether* it is activated at all and *how long* it remains activated depends on its current and subsequent motivational relevance to the organism. In other words, per ROAR, the organism has "relevance control" over its goals as it does over other activated knowledge, whether it is aware of it or not.

We conclude with a note on stability ("consistency") of behavior. We have documented the existence of representations that have cross-situational value relevance (faces; Eitam et al. 2013) and showed that these will operate regardless of a person's current task. Nonetheless, many representations have only transient relevance. Given that relevance determines activation, chronic relevance will lead to chronic activation and hence similar cognitions and actions across situations (consistency). Representations with transient relevance will make their appearance "inconsistently" or, rather, in a manner that is highly dependant on their current relevance to the situation at hand.

To summarize, although H&B highlight the central role of motivation in cognition and action, whether conscious or not, the chosen metaphor, that is, selfish goal, may unnecessarily limit their case in terms of capturing the scope of motivational effects on action and cognition and the degree of integration between cognitive and motivational influences. Unpacking their metaphor in terms of its underlying motivational-cognitive mechanisms could further strengthen its dominion.

Development links psychological causes to evolutionary explanations

doi:10.1017/S0140525X1300201X

Mark Fedyk[a] and Tamar Kushnir[b]

[a]*Department of Philosophy, Mount Allison University, Sackville, NB E4L 1G9 Canada;* [b]*Department of Human Development, Cornell University, Ithaca, NY 14853.*

mfedyk@mta.ca
http://markfedyk.net
tk397@cornell.edu
http://www.human.cornell.edu/bio.cfm?netid=tk397

Abstract: Our conscious abilities are learned in environments that have evolved to support them. This insight provides an alternative way of framing Huang & Bargh's (H&B's) provocative hypothesis. To understand the conflict between unconscious goals and consciousness, we can study the emergence of conscious thought and control in childhood. These developmental processes are also central to the best available current evolutionary theories.

Huang & Bargh (H&B) draw an analogy between the role that "selfish" genes play in Dawkins's account of evolution and "selfish," unconscious goals play in organizing behavior, and on this basis, propose a novel hypothesis about the proximate properties of unconscious goals. We argue that H&B's analogy risks the misleading impression that there is some kind of positive evolutionary evidence for their proximate hypothesis. Nevertheless, we find the proximate hypothesis is plausible in its own right and suggest an alternative way that its evolutionary significance can be defended without relying on the problematic analogy. Our alternative rests on two related ideas: First, to better understand the conflict between unconscious goals and conscious states, we can study the emergence of conscious thought and effortful control in childhood. This places more emphasis on developmental processes rather than on teleological metaphors. Second, developmental processes are central to the best available current evolutionary theories – in particular, theories that link evolution across multiple levels. H&B's admirable attempt to situate complex human behaviors into an evolutionary framework would be better served by these newer ideas.

H&B use Dawkins's "selfish-gene" metaphor to highlight their views about the functional organization of a person's unconscious goals. They also use the metaphor to ground their original proximate hypothesis, namely, that (in humans at least) goals are discrete mental states that give themselves their own "ends." But the only thing that H&B's goals and Dawkins's metaphorical view have in common is that both have their ends immanently. That is, for H&B, at least some unconscious goals are not given their ends from exogenous sources, like a higher-order conscious self, other goals, or environmental states. Yet, the commonalities between the two theories end with this structural similarity. Once an end is achieved, a goal does not necessarily recreate motivation for the same end, whereas no matter how many times they replicate, genes continue to "want" to do so by Dawkins's lights. Furthermore, unlike genes, there are no biochemical laws that specify when and under what conditions goals will be active; goals do not lay dormant for generations and only become active when "paired" with other goals. Genes do not compete with one another to produce developmental inconsistencies analogous to the behavioral "inconsistencies" that H&B find so striking. On these grounds, the analogy with Dawkins's is more misleading than helpful.

But there is a deeper problem. Resting their theory on this analogy risks making it appear that there is an argument from evolutionary premises to a conclusion stating the unconscious goal theory. There can be no such argument. This follows from the familiar distinction between proximate and ultimate explanations (Alcock 2001). Proximate explanations refer to organisms' "internal" properties, structural, and causal pathways that obtain during the life span of an organism, and they largely exhaust the types of explanations used in psychology, physiology, biochemistry, and neuroscience. By contrast, ultimate explanations refer to the environmental "external" factors, the causal influence of which is spread over generations, and as a result ultimate explanations are central to evolutionary biology and other historically focused sciences.

Proximate and ultimate explanations stand in a many-to-many relationship. Knowing the correct ultimate explanation for a trait or behaviour does not normally yield any meaningful inferences about which proximate explanations of the behaviour is correct; and the same is true in the other direction (Fedyk 2014). Here is the intuitive case for this: imagine a pattern of behaviour that is adaptive for some species in some environment, like cooperating with conspecifics. Then list several different proximate causes of the pattern of behaviour, like inherent prosociality or reputation management. Natural selection will not care which of these causes is the right one, and so generally speaking there is no inference from the fact that a pattern of behaviour is adaptive to any particular proximate conclusion. This is what explains the methodological separation of the psychological and biological sciences. More importantly, it also means that it is not possible to argue from ultimate explanations (e.g., genes replicate across generations) to proximate conclusions (e.g., goals are immanently selfish).

There is, however, much to be gained by exploring the evolutionary significance of unconscious goals. There is also much to be gained by better understanding why such goals persist in the face of our conscious abilities to metalize, plan, and deliberate, as well as our conscious skills (e.g., executive functions). To find proximate causes for H&B's most interesting examples of the conflict, we should trace its emergence, as unconscious goals of children begin to conflict with their developing abilities for conscious thought, effortful control, and self-regulation. Research on children's understanding of agency and intentionality, including their emerging beliefs about actions as free, autonomous choices, is especially relevant to H&B's aims (Kushnir 2012). This research helps us understand the influence of conscious goals on an organism that has already developed many unconscious goals. Importantly, this research can be undertaken without relying upon ultimate explanations.

That is not to say that these proximate explanations, once confirmed, cannot be linked up with equally well-confirmed ultimate hypotheses. If proximate causes for many of the phenomena described in this target article are better understood in developmental terms, then they are also better framed by more recent accounts that link evolution across multiple levels (genetic through social) and, as a result, have development as a central explanatory category (Jablonka & Lamb 2005; West-Eberhard 2003) . These theories use developmental explanations to emphasize the interaction of organisms with their environments. Concepts such as plasticity and niche construction become central. These ideas comport with our understanding of emerging consciousness in children, a process we know to be incredibly environmentally sensitive. Children learn about their minds and the minds of others through conversation, social interaction, and active cultural participation. In other words, our conscious abilities and skills are learned in environments that have evolved to support them.

The motivational self is more than the sum of its goals

doi:10.1017/S0140525X13002021

Ayelet Fishbach
Booth School of Business, University of Chicago, Chicago, IL 60637.
ayelet.fishbach@chicagobooth.edu
http://faculty.chicagobooth.edu/ayelet.fishbach/research/

Abstract: I present evidence in favor of an overarching motivational self: a mental function that regulates expression of multiple goals. Goals often conflict with each other, and the role of a motivational self is to consciously or unconsciously prioritize pursuit of these goals. When observing inconsistency in expression of goals, it is therefore useful to consider whether the motivational self is attempting to balance between conflicting goals or if such inconsistency results from temporary self-control weakness.

In explaining the interplay between multiple goals, one theoretical approach assumes an overarching entity that organizes expression of different goals: the "motivational self." In contrast, another approach assumes an unorganized and unguided competition between goals of different importance (e.g., goal shielding theory, Shah et al. 2002; the planner-doer model, Thaler & Shefrin 1981). The "Selfish Goal" model fits within the latter approach. It assumes goals operate autonomously in a somewhat random and chaotic order. The metaphor of genes competing for expression leads to the thought-provoking conclusion that there is no organization, coordination, or for that matter, a motivational self.

Huang & Bargh (H&B) present compelling evidence in favor of a Selfish Goal model. However, it might be also useful to consider the evidence in favor of a motivational self – that is, a centralized self-regulatory function that coordinates multiple goal pursuit, often outside of conscious awareness. Let me clarify that indeed people hold coexisting, often conflicting goals, and these goals compete with each other for limited resources (i.e., time, effort, and attention). The multiple goals people hold are also often inconsistent with each other (e.g., work and family goals) or directly undermine each other (e.g., saving and spending goals). I further agree with the assumption that situational cues (i.e., primes) trigger the activation of goals and influence judgment, feelings, and behavior. Thus, goal expression is subject to contextual cues, and because multiple goals coexist and people respond to different contextual cues, their goal pursuits often appear erratic.

However, there is evidence pointing at a motivational self that guides regulation of multiple goals. In what follows, I present three sources of evidence in favor of a motivational (centralized) self: the quest for multifinal means, the order of goal expression, and the exercise of self-control.

1. The quest for multifinal means. Goal Systems Theory suggests a guiding principle in action selection is the desire to find multifinal means, defined as means that serve multiple goals simultaneously (Kruglanski et al. 2002). For example, the task of choosing lunch food is often guided by (and a compromise between) several conflicting goals (e.g., hedonic pleasure, saving time and money, etc. Köpetz et al. 2011). Thus, individuals do not simply juggle between their goals. Rather, the motivational self searches for a compromise all goals can "live with."

2. Order of goals expression. Research on the expression of multiple goals suggests that what seems erratic is often orderly. There are two basic patterns of coordination between multiple goals: expression of one goal can increase the likelihood of expressing this goal again and inhibits competing goals, or expression of one goal leads to expression of another goal. The first pattern – "highlighting" goals – promotes consistency, whereas the second pattern – "balancing" between goals – promotes inconsistency. Indeed, consistency theories (Bem 1972; Festinger 1957; see also Bandura 1991) document a pattern of highlighting goals, whereas research on cybernetic self-regulation (Carver & Scheier 1998; Higgins 1987) and licensing (Monin & Miller 2001) documents a pattern of balancing goals (e.g., discriminating after expressing egalitarian views).

Research on Dynamics of Self-Regulation theory explores some of the variables that lead people to follow one order versus another. This research documented, for example, that uncommitted individuals and novices tend to highlight goal pursuit – they are more likely to adhere to a goal after they have pursued it. In contrast, committed individuals and experts balance their goals – they are more likely to adhere to a goal if they have previously pursued a different goal, in a dynamic of balancing (Fishbach et al. 2010). Research on learning goals makes a similar distinction: people respond to success by either increasing or decreasing academic pursuits, depending on people's implicit theory (Dweck & Leggett 1988).

Exploring the psychological variables that lead people to follow one pattern versus the other suggests an order in multiple goal pursuit, which implies a centralized motivational self. What appears to be inconsistency in goal expression often reflects a particular pattern: highlighting or balancing goals. And when researchers observe inconsistency, it is useful to ask if the goals were selected with respect to each other and if it is the pursuit of one goal that justifies pursuing another, inconsistent goal – for example, whether saving justifies subsequent spending behavior.

3. Self-control. Self-control research provides another source of evidence for a motivational self. According to self-control theory, individuals are not passive or helpless in pursuing multiple goals. These goals are often of varied importance: an individual would like to be financially responsible but is tempted to splurge, or she wishes to maintain good health while also feeling tempted by unhealthy foods. In these situations, self-control facilitates the pursuit of the more important goal and inhibit the tempting goal, often without conscious awareness (Fishbach & Converse 2010). Self-control is a limited resource (Baumeister et al. 1998) and people vary in their self-control ability (Mischel et al. 1989). Both temporary and personality weaknesses can explain giving in to temptation, in particular after a period of overriding goal pursuit. Thus, the motivational self's limitations, rather than its absence, often accounts for the pursuit of conflicting goals. When the motivational self is strong, behavior is often more orderly.

In summary, what appears as inconsistency often reflects a certain pattern of goal expression and coordination. There is evidence for a guiding, overarching self-regulatory system, which organizes multiple goal pursuit: the motivational self. It sometimes fails but nonetheless exists, and it is more than the sum of one's goals.

Mapping the goal space: Personality integration and higher-order goals

doi:10.1017/S0140525X13002033

Jacob B. Hirsh
University of Toronto, Toronto, ON M5S 3E6, Canada.
jacob.hirsh@utoronto.ca
http://www.jacobhirsh.com

Abstract: By situating goals at the heart of human cognitive function, Huang & Bargh (H&B) provide a useful platform for understanding the process of personality integration as the gradual mapping of implicit motives into a coherently organized self-system. This integrative process is a critical feature of human development that must be accounted for by any complete goal theory.

In their target article, Huang & Bargh (H&B) describe goals as autonomously operating motivational structures that direct human cognition and behavior. Such goals often operate unconsciously, shaping an individual's experience of the world without any explicit awareness. Importantly, these goals are argued to function in a "selfish" manner, *as if* each goal structure was concerned only with its own completion. The present commentary examines the implications of Selfish Goal Theory for models of personality integration and highlights the importance of the integrative process for H&B's framework.

Personality integration, reflecting the extent to which the various aspects of one's psyche function as an integrated whole, has long been regarded as a core process in psychodynamic and humanistic theories of development (Allport 1937; Jung 1939; Maslow 1970; Rogers 1951). The developmental process of recognizing one's diverse motives and incorporating them into a fully elaborated self-concept has been described variously as individuation, self-actualization, self-realization, or self-discovery. More recent formulations of personality integration define it as the extent to which an individual's explicit goals are aligned with each other and with basic human needs (Emmons 1986; Sheldon & Kasser 1995). The integrative process of identifying and aligning one's goals is important to understand because it is directly related to psychological well-being: greater personality integration and goal alignment can reduce conflict-related stress while increasing vitality, motivation, and personal meaning (Emmons & King 1988; Hirsh et al. 2012; McGregor & Little 1998).

Integrating personal goals into a coherent self-system can require a great deal of effort and self-reflection. Because implicit motives are often pursued in the absence of explicit goal representations, many people have only a partial understanding of the goals that shape their behavior (Bargh & Barndollar 1996). During development, children only gradually come to a conscious understanding of their implicit motives as the self-concept is elaborated over time (Damon & Hart 1982). Prior to the explicit recognition of these desires, an individual's motives will be pursued largely unconsciously (Hoffree & Winkielman 2012). Even among adults, the elaboration of implicit goals into a fully articulated self-understanding can be a lifelong learning process. Within any given person, some goals will have been clearly articulated and incorporated into the self-concept, while others will remain outside the realm of conscious awareness (where they nonetheless have a strong impact, as H&B argue).

By emphasizing the central role that goal constructs play in shaping human thought and action, Selfish Goal Theory provides a useful platform for modeling the process of personality integration and self-development. In particular, this process can be fruitfully understood as the gradual mapping of implicit goal space and the alignment of these goals within the self-system. An individual thus begins his or her journey to personhood as a jumble of motivational impulses that direct cognition and behavior with minimal self-awareness. Such a person would clearly demonstrate H&B's "selfish" goal effects, including behavioral inconsistency and poor self-insight, amid the multitude of implicit goals vying for control of the behavioral system. As the process of self-development and personality integration unfolds, however, these implicit goals will eventually be elaborated into explicit representations that are incorporated into the self-concept. This developmental process is equivalent to the psychodynamic integration of the unconscious mind into the conscious mind (Jung 1939) and the humanistic striving for authenticity as reflected in congruence between organismic needs and conscious behavior (Rogers 1951).

An important implication of this process is that individual goals will become less "selfish" after being integrated within a larger self-system. In its current form, Selfish Goal Theory focuses on the competition between autonomous goals, each of which operates as a functionally segregated behavioral control system. Although a completely unintegrated person might indeed shift from one salient goal to the next with no behavioral consistency, the coordination of human goals can be characterized by a much greater degree of order and self-regulation. In particular, the process of personality integration binds an individual's disparate motives into a single integrated goal structure, such that the engagement of various goals can be regulated in the service of long-term plans and higher-order values (Austin & Vancouver 1996; Carver & Scheier 1998; Peterson 1999). These higher-order goal systems can influence the relative salience of different lower-order goals and are thus critical for understanding the self-regulation of human thought and behavior (Kruglanski et al. 2002). For example, the distinct goals of being productive at work and sticking to a fitness plan can both be integrated within the higher-order goal of being a successful person. The ability to effectively regulate basic goals within higher-order goal systems is in fact one of the distinguishing characteristics of the human nervous system, as instantiated by the integrative functions of the anterior prefrontal cortex (Hirsh 2010; Koechlin et al. 1999).

Individuals with a more fully elaborated self-understanding and explicitly articulated value system are thus more likely to experience identity coherence and behavioral consistency, owing to their more deeply integrated goal structures. By incorporating basic motives within higher-order goal systems, more integrated individuals will also alleviate the stress and anxiety that arises from conflicts between otherwise competing goals (Emmons & King 1988; Hirsh et al. 2012; Nash et al. 2011).

Selfish Goal Theory presents an intriguing framework for understanding the tremendous power that goals have in shaping cognitive-behavioral dynamics. In order for it to provide a complete account of goals and their substantial cognitive impact, however, the theory could benefit by more directly addressing the role of integrative processes and higher-order goal systems in the global organization and self-regulation of human behavior. In its current form, the theory appears to be most effective at describing goal dynamics in the limiting case in which a person's goals are functionally segregated from one another and unrepresented by the self-system. As these goals become explicitly linked within a hierarchical system for behavioral control during the process of personality integration, the competition between multiple "selfish" goals may no longer provide a complete description of human goal dynamics.

Massively representational minds are not always driven by goals, conscious or otherwise

doi:10.1017/S0140525X13002045

Bryce Huebner[a] and Robert D. Rupert[b]

[a]*Department of Philosophy, Georgetown University, Washington, DC 20067;*
[b]*Department of Philosophy and Institute of Cognitive Science, University of Colorado, Boulder, Boulder, CO 80309-0232.*
lbh24@georgetown.edu
http://www9.georgetown.edu/faculty/lbh24/
robert.rupert@colorado.edu
http://www.colorado.edu/philosophy/fac_rupert.shtml

Abstract: The language of conscious and unconscious goals is rooted in a folk-taxonomy that is likely to inhibit progress in cognitive science. Severing the commitment to this taxonomy would allow Huang & Bargh (H&B) to consider a wider variety of representational forms with motivational force and to entertain the intriguing possibility that variations in the number of active-but-redundant representations account for variance in social behavior.

Social psychologists often avoid questions about cognitive mechanism – and for good reason. Socially significant behavior is complicated, and it is fiendishly difficult to pull apart the relative contribution of environmental factors, representational states, and computational processes. The hypothesis that variations in behavior can be "meaningfully understood as the output of multiple, and in some cases, competing goal influences" (sect. 1, para. 2) promises to advance our understanding of the mechanistic underpinnings of social cognition. But understanding these "goal influences" in their full generality requires looking beyond winner-take-all competitions between "selfish goals" and "representations of desired end-states" (sect. 2.2, para. 2). We encourage Huang & Bargh (H&B) to follow their logic to a natural, if more radical, conclusion – to recognize that various forms of representation have motivational force and that the language of conscious and unconscious goals is rooted in a taxonomy likely to inhibit progress in cognitive science. We embrace the thought that numerous, competing influences run in parallel and collectively affect behavior, but this thought is best pursued without the constraints imposed by the folk notion of goals, Dawkinsian metaphors of selfishness, and the assumption that conscious goal pursuit is particularly distinctive.

There is growing evidence that human cognition relies on numerous, sometimes redundant, sometimes overlapping representational systems, and that it trades in multiple representations of the same thing (e.g., properties, situations, objects; cf, Huebner, 2013; Rupert 2011). Some of these representations are abstract and linguistic, others are action-oriented representations (Clark 1997; Millikan 1995), and still others appear as parts of emulator-circuits and forward-models (Grush 2004; Shultz et al. 1997). All of these types of representation seem to contribute to the production of goal-directed behavior, and their heterogeneous nature suggests that it is unwise to treat representations of end-states as the mind's ultimate motivational foundation.

Consider the representation "do *A*, if condition *C* is met." Neither *A* nor *C* is an end-state, and it is unclear whether this representation counts as a goal (or whether it matters). H&B (sect. 4.3) rightly note that a person will recognize that *C* is met more quickly in tasks that are associated with *C*, and they take this fact to support *reconfiguration*. Specifically, they claim that the desire to follow instructions primes a subject to recognize *C*, thereby facilitating the recognition of *C*. But in a parallel processing system that utilizes competition, it is equally likely that the prime-related words appearing in the instructions activate more representations of *C* than would be active at baseline (cf Gendler 2008a; 2008b; Schröder & Thagard 2013), that visual stimuli activate additional representations of *C*, and that this parallel activity pushes the cognitive system past the threshold for *A*-production more quickly than would visual processing on its own. This does not appear to be a process by which a representation of an end-state commandeers cognitive resources. In general, fast and accurate biological motion is likely to require the integration of multiple representations from emulation-circuits, as well as perceptual and goal-driven systems; thus, manipulating an object in a way that violates sensory expectations generates conflicts between visual and motor representations, resulting in incompatible task-demands, as well as slower behavior, even where there are no representation of end-states, or of such goals as pleasing an experimenter (Wexler et al. 1998).

We contend that human behavior is often affected by the accumulation of redundant representations, some of which are produced by independent subsystems and some of which are produced within the same subsystem. Admittedly, this talk of multiple, redundant representations is likely to sound strange, and this is part of our point. The claim that *goals* reconfigure *cognitive* states presupposes a folk-psychological taxonomy, but the operation of *cognitive mechanisms* often outstrips the capacities readily described in such terms. This fact becomes especially clear in cases where competitive algorithms integrate, strengthen, inhibit, and recode the various representations distributed throughout a cognitive system (Akins 1996). By severing the commitment to traditional folk-taxonomies, we can examine the theoretically and computationally important possibility that variations in the number of active-but-redundant representations might account for variance in reaction times and error rates.

Traditional assumptions about psychological processing also make conceptual mischief in the discussion of *similarity*. H&B (sect. 4.4) embrace standard assumptions about our awareness of goals, our conscious intentions to pursue them, and our experience of control over the goals that we pursue. Thus, they find it striking that both conscious and unconscious behavior manifest deep similarities. Yet, they also acknowledge that the conscious self "is not so much involved in the guidance of our purposive behavior so much as it is in the business of producing rationalizations and socially acceptable accounts for the actions produced at the goal level" (sect. 5.2, para. 9). This latter view yields *similarity* almost for free. As H&B note, conscious goal pursuit relies on the mechanisms that guide unconscious goal pursuit (conscious states are "built upon preexisting unconscious processes" [sect. 1, para. 5]), with the addition of mechanisms for rationalization. But if *goal pursuit* is implemented by the same mechanisms – regardless of whether the subject can report accurately on her motivation – *similarity* holds trivially. Admittedly, conscious goal pursuit might engage working memory and meta-cognition in distinctive ways, but, primarily, this reflects the activity of a processing stream that independently produces representations grounding verbal report, rationalization, and the experience of conscious will. Sometimes these representations match the content of the processes that produce behavior (and *sometimes* this yields facilitation); sometimes they do not (and *sometimes* this yields inhibition); and sometimes conscious representations are absent altogether.

It is at least a rhetorical mistake to distinguish two fundamental forms of goal pursuit – conscious and unconscious – and to note their similarities; better to think of a single, perhaps complex, form of motivational processing, and to ask how motivational processes interact with systems that produce verbal report. Although we agree that numerous motivational states collectively affect behavior, we think that H&B could strengthen their position by adopting a broader taxonomy of state-types. By abandoning the emphasis on (selfish) goals and the distinction between their conscious and unconscious forms, H&B might pursue a massively representational model of mind that fully harnesses the resources afforded by the cognitive sciences.

Theoretical integration in motivational science: System justification as one of many "autonomous motivational structures"

doi:10.1017/S0140525X13002057

Aaron C. Kay[a] and John T. Jost[b]
[a]*Fuqua School of Business and Department of Psychology & Neuroscience, Duke University, Durham, NC 27708;* [b]*Department of Psychology, New York University, New York, NY 10012.*
aaron.kay@duke.edu
https://faculty.fuqua.duke.edu/~ack23/index.html
john.jost@nyu.edu
http://www.psych.nyu.edu/jost/

Abstract: Recognizing that there is a multiplicity of motives – and that the accessibility and strength of each one varies chronically and temporarily – is essential if motivational scientists are to achieve genuine theoretical and empirical integration. We agree that system justification is a case of nonconscious goal pursuit and discuss implications of the fact that it conflicts with many other psychological goals.

Social psychologists have gained considerable traction in studying the epistemic, existential, and relational underpinnings of attitudes and behavior (Greenberg et al. 2004; Jost et al. 2009). Diverse, seemingly incongruous programs of research have helped to illuminate the structure and function of human needs and the causes of seemingly paradoxical, self-defeating behavior – why, for example, "the freed bird finds a new cage" (Pyszczynski et al. 1995).

Unfortunately, researchers are often drawn into unhelpful debates about which single theory or motive – for example, social identification, terror management, system justification, social dominance, or meaning maintenance – subsumes the others or is otherwise more "basic" or "fundamental." The problem is that competition among "rival" theories tends to be illusory, and such debates are rarely, if ever, resolved (Greenwald 2004; Sullivan et al. 2012). Another, quite different approach is to distill a common denominator (such as inconsistency compensation) and conclude that seemingly disparate theories are really all saying the same thing (Proulx et al. 2012). We share the scientific values of parsimony and unification, but this approach is also unsatisfying, insofar as it obscures important differences in theoretical insight and diminishes our ability to distinguish between (sometimes vastly) different behavioral outcomes.

Far more promising is Huang & Bargh's (H&B's) proposal that the motivational system is itself composed of multiple, potentially conflicting goals that vary chronically and temporarily in terms of activation and strength. We share the authors' supposition that system justification – defined as the motivation to defend, bolster, and justify aspects of the societal status quo – is one of many "autonomously operating motivational structures." The basic idea is that people are generally unaware that they possess the goal of imbuing social, economic, and political systems with legitimacy – but they do so anyway (Jost et al. 2008).

Evidence is accumulating that system justification operates as a largely nonconscious goal. For example, we know that exposure to criticism of the system stimulates defensive motivation (Kay et al. 2009), as well as automatic positive evaluation of societal symbols and heightened accessibility of concepts related to legitimacy and social stability (Liviatan & Jost 2014). These effects persist only as long as the system justification goal is active; they tend to disappear when the goal is satiated through acts of direct system-affirmation. Otherwise, system justification persists even in the face of obstacles, so that individuals work hard (even on impossible tasks) in an effort to affirm the legitimacy of "the American dream" (Ledgerwood et al. 2011).

At the same time, system justification frequently conflicts with other goals that the individual possesses, including goals to maintain self-esteem (ego justification) and to defend the interests or actions of fellow group members (group justification). Members of disadvantaged groups, such as women and minorities, appear to suffer from motivational conflicts involving ego, group, and system justification concerns. Indeed, system justification is a perfect example of what H&B refer to as a "self-defeating" goal. For example, African Americans who justify the status quo possess lower levels of self-esteem and higher levels of depression and neuroticism (Jost & Thompson 2000), and women who justify the status quo exhibit a depressed sense of entitlement, coming to believe that they deserve inferior wages (McCoy & Major 2007; O'Brien et al. 2012).

The point of system justification theory is not to suggest that everyone *always* justifies the societal status quo (cf. Reynolds, et al., in press) – or that those who are disadvantaged are necessarily the ones who justify it the most (cf. Brandt 2013). Often, the goal to defend or bolster personal or collective self-esteem is stronger or more active than the goal to defend or bolster the social system. The point is that nearly everyone – including members of disadvantaged as well as advantaged groups – possesses *some* (variable) degree of motivation to believe that the systems on which he or she depends are fair, legitimate, and so on. There is no reason to assume that people will *always* justify the status quo because they may also possess other goals (in addition to ego and group justification), such as goals to be egalitarian, fair, accurate, innovative, contrarian, or vengeful.

This has been illustrated vividly in experiments demonstrating that people respond to the same social or political stimulus (such as proposed legislation that infringes upon personal freedom) by either rationalizing (Kay et al. 2002) or reacting against it (Brehm 1966). Laurin et al. (2012) demonstrated that reactance and rationalization are *both* motivational structures that operate autonomously in different sets of circumstances. Individuals strive to resist infringements upon freedom of choice when the possibility remains that the new restrictions will not be enacted. When the same infringements are described as inevitable, however, individuals tend to accept and begin justifying the new regime.

The fact that human beings are capable of divergent responses to the same stimulus event makes evolutionary sense. It may be that investing resources to resist external forces (such as the impositions of social or political systems) is adaptive only when there is some chance, however remote, that the impositions can be avoided. When this is impossible, it may be more adaptive to accommodate new realities and focus on "silver linings" rather than committing resources to a fight that appears to be unwinnable. This is a useful, integrative conclusion that would have been excluded from typical forms of theory competition, such as the search for a "critical test."

Rather than debating, perhaps indefinitely, which motive is the strongest or most fundamental for human behavior, we believe that researchers would be better off using their prodigious theoretical and methodological skills to determine how and why a given goal-state operates autonomously and when it is especially likely to guide behavior. We applaud H&B for helping to show us the way forward. Recognizing that there is a multiplicity of (sometimes conflicting) motives – and that each one varies in terms of accessibility and strength as a function of the person and the situation – is essential if we are to achieve genuine theoretical and empirical integration in motivational science.

On the selection and balancing of multiple selfish goals

doi:10.1017/S0140525X13002069

Catalina Kopetz,[a] Wilhelm Hofmann,[b] and
Reinout W. H. J. Wiers[c]

[a]*Department of Psychology, Wayne State University, Detroit, MI 48202;* [b]*University of Cologne, 50931 Cologne, Germany;* [c]*Department of Psychology, University of Amsterdam, 1018 XA Amsterdam, The Netherlands.*

catalina.kopetz@wayne.edu
http://clasweb.clas.wayne.edu/catalina-kopetz
http://www.addiction.umd.edu/CAPERWebSite/CatalinaKopetz.html
r.w.h.j.wiers@uva.nl
http://www.uva.nl/over-de-uva/organisatie/medewerkers/content/w/i/r.w.h.j.wiers/r.w.h.j.wiers.html
wilhelm.hofmann@uni-koeln.de
http://soccco.uni-koeln.de/wilhelm-hofmann.html

Abstract: The selfish goal metaphor is interesting and intriguing. It accounts for the idiosyncrasies and inconsistencies in peoples' goal pursuits without invoking free will, self-regulatory, or self-control failures. However, people pursue multiple goals, sometimes simultaneously. We argue that the model proposed in the target article may gain significant theoretical and practical value if the principles underlying goal selection and/or balancing on a moment-to-moment basis are clearly specified and integrated with the notion of the selfish goal.

Huang & Bargh (H&B) propose a metaphor whereby, just like genes, goals operate autonomously according to their own agenda even when doing so may not be in the overall interest of their host/individual. Although this metaphor is interesting and intriguing, we believe that the story of goal pursuit is more complicated than one goal selfishly pursuing its agenda at the expense of other goals. Goals do not come in isolation, but rather in configurations of multiple goals. They may become simultaneously active and may facilitate or inhibit each other, with different consequences for action, as well as for individuals' phenomenological experience (Kruglanski et al. 2002). Once a goal is selected, it may pursue its agenda selfishly. However, H&B's metaphor may result in an overly simplistic portrayal of human action without (1) specifying the principles and consequences (i.e., the experience of goal conflict) of goal selection and (2) acknowledging that sometimes people may pursue multiple goals simultaneously whereby no goal selection is needed.

Regarding the issue of goal selection, H&B argue that each and every goal may pursue its agenda selfishly. But, given that individuals have a multitude of goals that are often activated simultaneously, it is important to understand the principles through which a particular goal takes priority and gets to pursue its agenda. Although the criterion is very clear for genetic selection (i.e., survival), it is not as clear for goal selection. H&B propose that when more than one goal is currently active, "the most incentivized" (sect. 2.1, para. 5) of these wins. However, it remains unclear (1) how "incentivization" is defined and operationalized, (2) and what mechanisms are responsible for selecting the most incentivized goal. Without clarifying these issues, the argument becomes almost circular (the winning goal must have been the most incentivized one) and specific predictions about an individual's behavior can be neither made nor tested.

H&B do talk about unconscious inter-goal inhibition and introduce "the reconfiguration principle" (sect. 4.1, para. 1) to explain goal selection. However, such selection seems to be accounted for solely by goal activation. "The most motivating or 'active' goal should constrain … behavioral possibilities in a way that encourages achievement of the goal's end-state" (sect. 4.1, para.1). However, goal pursuit is not only a function of goal activation. The value attached to that particular end-state, as well as the expectancy of attaining it, determines goal-directed behavior, in addition to goal activation. More important, both the value and the expectancy of goal attainment, on a moment-to-moment basis, are relative to the other goals that may be currently active (Atkinson 1964). For example, an active goal of smoking may lose its value and may not result in smoking behavior when the negative consequences of smoking (and presumably a health-related goal) become active. (Kober et al. 2010). Active goals may inhibit competing alternatives resulting in goal shielding (as acknowledged by H&B). But goal shielding is a function of goal value (i.e., importance), not only of goal activation (Shah et al. 2002). Furthermore, goal shielding is not a necessary consequence of having multiple goals activated simultaneously. Indeed, goals may sometimes facilitate each other resulting in goal co-activation, rather than goal shielding (Shah et al. 2002.). This raises the intriguing issue that, whereas goals may appear selfish when viewed in isolation, they may under certain conditions benefit associated, but distinct goals (e.g., the goal to eat healthily could activate the goal to exercise). The selfish goal account appears to neglect the possibility of mutually supporting goal architectures.

Secondly, the selfish goal metaphor fails to predict behavior in situations when multiple goals of equal importance are activated. In such situations, people will strive to pursue them simultaneously by finding multifinal means (Kopetz et al. 2011; Kruglanski et al. 2013). This may explain why someone who is hungry and therefore has the goal of eating, but who is also on a diet (pursues the goal of weight watching) may go for a salad rather than a double cheeseburger with fries. However, when the expectancy of finding multifinal means is low, individuals will engage in goal selection and pursue the goal of highest value (Kopetz et al. 2011).

Finally, the selfish gene approach has been criticized for failure to consider higher levels of selection (i.e., group level) (Goodwin 2001). An analogous criticism may apply to the notion of selfish goals. Goals may selfishly pursue their agenda at the expense of other goals on a moment-to-moment basis. Furthermore, conscious choice or free will does not need to be invoked to explain behavior. However, such selfish pursuit may result (post hoc) in experienced goal conflict, which has known negative consequences (Emmons & King 1988). Such negative consequences, whether experienced or anticipated, may prompt people to develop strategies to avoid goal conflict (Fishbach et al. 2009; Kopetz et al. 2011). Individuals may differ with regard to how they balance, integrate, and make sense of their multiple selfish goals. Individuals with more balanced goals may have a selective advantage because they advance their goals in a better-coordinated manner on multiple domains in life (Bélanger et al. 2013). In contrast, a poorly concerted set of selfish goals may lead to psychopathology. From this perspective, the metaphor of the selfish goal may be misleading and shortsighted insofar as it does not concern itself with aspects of higher-order integration.

To summarize, we believe that the selfish goal metaphor is interesting and intriguing. It acknowledges that people do what they do for a reason (an active goal) even if the momentary behavior may appear in contradiction with the persons' long-term best interest. In that, it explains many aspects of goal pursuit and accounts for the idiosyncrasies and inconsistencies in people's behavior without invoking free will, self-regulatory or self-control failures. However, given the multitude of goals that people pursue, sometimes simultaneously, the model may gain significant theoretical and practical value if the principles that account for goal selection and/or balancing on a moment-to moment basis are clearly specified and integrated with the notion of the selfish goal.

The validity of Dawkins's selfish gene theory and the role of the unconscious in decision making

doi:10.1017/S0140525X13002070

Tobias A. Mattei
Neurosurgery Department, Ohio State University, Columbus, OH 43201.
tobias.mattei@osumc.edu
http://www.invisionhealth.com/providers/neurosurgery/tobias-mattei-md/

Abstract: Although the proposed Selfish Goal Theory constitutes a major theoretical tour de force for addressing the issue of inconsistencies in human actions and the role of motivational goals in behavior, as it is based on an unproven biological paradigm (Dawkins's selfish gene theory) and overemphasizes the role of unconscious processes in decision making, it provides a questionable model of the underlying psychological structure of human agency.

Huang & Bargh (H&B) propose a theory for interpretation of human behavior based on a psychological analogy of Dawkins's selfish gene paradigm. According to such framework, instead of a unitary and coherent multitasking coordination apparatus, human beings would operate under the continuous (and, in most cases, unconscious) influence of a dynamic selection process among multiple task-generating goals, which would be constantly competing for psychological resources from different cognitive faculties such as attention, emotional salience, language, and analytical reasoning.

Although H&B assume the empirical validity of Dawkins's selfish gene theory, several authorities in physiology and evolutionary biology (Huang 2012; Noble 2011) have considered such a proposal a simplistic, metaphoric, and untestable hypothesis that presents close similarities with old mythological interpretations of natural events. In fact, Dawkins himself recognized: "I doubt that there is any experiment that could prove my claim" (Dawkins 1982). Affirming an "apparent purposiveness of evolution" or stating that the "psychological and behavioral phenomena are designed by natural selection for a purpose" applies exclusive human attributes (such as intentionality) to purely biological processes. This naive intellectual modus operandus based on anthropomorphic analogies has been considered by several authors as a classic form of early pre-scientific reasoning (Agassi 1968; Dundzila 1987; Hordern 1972) in which objective natural processes are "personalized" and interpreted as functioning according to the same human framework that involves analytical reasoning, emotions, and purpose. Interestingly, Dawkins, as a neo-Darwinian, is not far from his master, who, although widely considered an epitome of a scientist also seems to have appreciated the use of anthropomorphic reasoning as illustrated by the following quote (Darwin 1871, pp. 39–40): "Even insects play together, as has been described by that excellent observer, P. Huber, who saw ants chasing and pretending to bite each other, like so many puppies."

Additionally, at the present time, the so-called "central dogma of molecular biology" proposed by Crick (1970) (which consists of a simplistic paradigm of "DNA→transcription→RNA→translation→protein") has been strongly questioned by new research findings in the emergent field of systems biology (Bor-Sen et al. 2011; Longo & Tendero 2007). Actually, it has already been demonstrated that several other mechanisms involving epigenetic regulation may play an important role in information encoding within the cell, such as the three-dimensional tertiary structure of proteins that can be changed by allosteric transitions (Shapiro 2009), and multitasking RNA molecules like riboswitches, microRNAs, and small interfering RNAs (Mirouze 2012).

Moreover, the human genome is more than 98% noncoding DNA (Pennisi 2012), which do not operate "selfishly" but serve as regulators of the transcription/translation process. Following such discovery, some began advocating an evolutionary theory focused on whole cellular subsystems, such as the aerobic respiratory or the photosynthetic system, because theories focusing on individual genes cannot account for the emergence and preservation of important portions of genome such as promoters, enhancers, operators, and transcription factors (Noble 2011). In the same way, if we adopt an individualistic view of genes as operating apart from their ultimate effects on cellular functions, as "semi-parasitic" entities whose major goal is to maximize their propagation into future generations despite the final outcomes to cellular homeostasis, it is very hard to explain the evolutionary emergence of apoptosis, cell-cycle regulation, and tumor suppressor genes (da Fonseca et al. 2010; Pearson & Sánchez 2008).

Besides Dawkins's original selfish gene theory, another major theoretical pillar of the proposed Selfish Goal model is the assumption that human decision making processes are the evolutionary product of a much older system, in which unconscious mechanisms directed by different environmental-generated goals elicit a psychological competition for the cognitive resources required to accomplish survival-related tasks. Since early pioneering studies in psychology (Freud 1915; James 1890; von Schelling 1800), unconscious processes have exerted a significant attraction to researchers and several contemporary experts have strongly emphasized the activity of an "unconscious mind" (Kihlstrom 1987; Schacter 1992; Wilson 2002). Nevertheless, in a recent evaluation about the scientific merit of such claims, Newell and Shanks (2014) severely criticized the methodology of previous experimental studies that were considered as evidence for the decisive role of unconscious factors in human choices. In fact, according to these authors, unconscious influences in decision making and behavior have acquired an inflated and erroneous explanatory power in current psychological theories that is not supported by experimental evidence.

In the proposed Selfish Goal Theory, consciousness figures predominantly as an epiphenomenon of the human mind, not providing any clear evolutionary advantage because H&B argues that most of the decision making is governed at the unconscious level. Such a paradigm goes against the basic assumptions of current neuroscience theories of self-awareness (Morin & Michaud 2007) and agency (Haggard & Clark 2003; Kühn et al. 2013), which strongly emphasize the evolutionary emergence of consciousness as a cognitive framework through which multisensory and intertemporal events can be coordinated through a highly elaborated unitary cognitive faculty (Cabanac et al. 2009; Donald 1995; Pennartz 2009). According to such a viewpoint, consciousness has emerged as the by-product of the activity of the multimodal integrative areas of the human brain that receive inputs from a diffuse neocortical network operating through hierarchical levels of information analysis. Such neural correlates of consciousness would be closely connected with the limbic system and, by providing an estimation of the self-relevancy of sensory stimuli (as well an overview of their relation to previous beliefs and future expectations), would lead to emergence of an unique first-person perspective experience of reality (Vogeley & Fink 2003) with the ultimate goal of supplying the premotor areas of the brain with coherent sequential motor plans.

In summary, because H&B's theory is based on a metaphoric and, at best, unproven biological paradigm (Dawkins's selfish gene theory) and because it puts an excessive emphasis on the role of unconscious processes in decision making (which is not warranted by experimental evidence), the proposed Selfish Goal model, although constituting a major theoretical tour de force for addressing the major issue of intertemporal inconsistencies in human behavior, provides a questionable interpretation of the underlying psychological structure of human agency.

The effects of being conscious: Looking for the right evidence

doi:10.1017/S0140525X13002082

Marco Mazzone
Department of Human Sciences, University of Catania, 95124 Catania, Italy.
mazzonem@unict.it
http://unict.academia.edu/MarcoMazzone

Abstract: Huang & Bargh's (H&B's) general picture might underestimate the role played by conscious self and overestimate the behavioral inconsistencies at the personal level. This follows from how they delimit the goals under consideration: Their theses that goals are not consciously selected and that the conscious self is involved just in post hoc rationalization should also be tested against concrete and long-term goals.

Because I find that in Huang & Bargh's (H&B's) target article there is much that is agreeable, my point is not *whether* their notion of selfish goal is correct and useful but rather *to what extent* it is. Another way to put this is to say that their article attempts to draw distinct conclusions that are not equally supported. Specifically, it is one thing to say that (1) "goals influence a person as if the goals themselves were selfish and interested only in their own completion" (abstract, para. 1), it is another to say that (2) "a person's behaviors are indirectly selected at the goal level but expressed (and comprehended) at the individual level" (abstract, para. 2), it is yet another to say that (3) the conscious self "is not so much involved in the guidance of our purposive behavior so much as it is in the business of producing rationalizations" (sect. 5.2, para. 9). The first conclusion is strongly supported by H&B's arguments, especially by what the authors dub the "reconfiguration principle": the evidence that goals constrain "the person's cognitive and affective machinery for the purpose of facilitating goal pursuit" (sect. 4.3, para. 1) irrespective of whether those goals are conscious or not. That conclusion, however, does not imply that goals are never selected at the individual – or, as I would say, the "personal" – level. And the claim that goals are not selected at the personal level does not imply that, once selected, conscious goals cannot have a role in guiding behavior over and beyond mere rationalizations.

The fact that those conclusions are not equally warranted is somehow concealed by H&B's choice to limit their considerations to a specific subset of the goals underlying actions. H&B make it explicit that they are "more interested in time-limited instrumental behaviors enacted in the current situation" (sect. 2.2., para. 2) than in long-term, chronic goals; besides, they are less focussed "on the very concrete actions that can be described in fully objective terms (e.g., pressing a button; opening a door)" (sect. 2.2., para. 2) than "on higher-level end-states that provide those actions with meaning" (sect. 2.2., para. 2). A typical example of the goals they are concerned with is the goal of being either cooperative or competitive in a given time-limited task.

Here, the first thing to note is that abstract goals of this sort are not the ones that are ordinarily expected to fall within the focus of attention. Any human action involves the representation of a multiplicity of goals at different levels of abstraction, not all of which can be consciously attended during processing (Mazzone & Campisi 2013). Because it is possible that we consciously attend concrete goals more than higher-level ones (e.g., the goal of performing a task more than the goal of being cooperative or competitive while performing that task), then by restricting considerations to the latter, H&B may have failed to look in the right place in order to assess the role of consciousness in guiding behavior.

To be sure, H&B insist that there are strong similarities between conscious and unconscious goals (they call it the "similarity principle"). But this can hardly mean that it makes no difference whether conscious processes are involved in guiding behavior or not. H&B implicitly acknowledge this point insofar as they recognize in passing that conscious processes are superior to unconscious ones "in serving integrative functions" (sect. 5.1, para. 1) and allow for "metacognitive abilities" and "tighter executive control over the initial impulses to action" (sect. 3.4, para. 2). Moreover, the cognitive system for guiding behavior is said to be "less centralized in the absence of consciousness" (sect. 3.1, para. 4). But then, contrary to the thesis (3) above, it seems not to be the case that conscious and unconscious goals – though similar in some respects – have identical functions except for the role that conscious goals play in post hoc rationalizations. The above quotations suggest a role for consciousness in the assessment of coherence between goals and, therefore, in their integration. This could have effects that go well beyond mere post hoc rationalizations for the purpose of social accountability.

Let us note that the notion of selfish goal is said to account for why there are inconsistencies in human behavior at the personal level. But again, one would wonder whether the role of conscious self in preventing inconsistencies is not underestimated as a consequence of how H&B select the goals under consideration. Not only, as noted above, abstract motivations, such as being cooperative, are less likely to be consciously attended than concrete goals. Moreover, abstract motivations of that sort are not the kind of things that can be selected once and for all. On the contrary, we expect that there is a balance between cooperation and competition to be struck on a case-to-case basis: people are not expected to be either coherently cooperative or coherently competitive across any possible action. In order to assess whether goals are coherently pursed or not, we should rather look to more concrete goals and especially to long-term goals. It is here that we can possibly find cases of conscious selection and the effects of conscious adoption of goals.

One case in point is that of addictions. H&B themselves repeatedly address this issue, but their emphasis is on automaticity. However, one might focus instead on how addictions can be resisted when the conscious self is engaged in a search for rational coherence. For example, people may decide to give up smoking because they appreciate the rational relationship with their larger purpose of pursuing wellness (goal selection). Moreover, the decision to give up smoking may affect behavior in the long run in many ways, including the conscious search for appropriate strategies to that end. Rational coherence has some role to play here.

In sum, conscious and unconscious goals similarly reconfigure the person's cognitive and affective machinery. Besides, goals can be automatically selected and pursued. But they can also be *consciously* selected and pursued, and this may increase coherence at the personal level.

Genes, hosts, goals: Disentangling causal dependencies

doi:10.1017/S0140525X13002094

Bjorn Merker
Kristianstad, Sweden SE-29194.
gyr694c@tninet.se

Abstract: The special sense in which the concept of "selfishness" is defined in Dawkins's popularization of basic evolutionary theory is analyzed with regard to its applicability to the relation between goals and those who entertain and pursue them. It is concluded that grounds analogous to those on which independent self-interest vis-á-vis their hosts is attributed to genes in Dawkins's sense are lacking in the case of goals in their relation to those who entertain and pursue them.

In their target article, Huang & Bargh (H&B) introduce and advocate the adoption of a novel explanatory construct in the domain of cognition and behavior, explicitly conceived by analogy to Richard Dawkins's "selfish gene." H&B support their analogy by the apparent parallelism between genes using organisms as propagation vehicles "sometimes to the detriment of the host organism's life" and human goals that, whether consciously held or not, can impair the lives of those who pursue them. Analogies have their uses in science, as do metaphors, provided that – in either case – they are apt and apposite. To know whether the similarity between genes and goals just mentioned betokens more than the commonplace that functional mechanisms have limitations and sometimes malfunction, we must inquire into the typical function of genes and goals – which cannot be to impair the lives of their hosts – and ask whether any grounds analogous to those on which selfishness is attributed to genes exist for goals as well.

Genes are not only typically, but *always* "interested" (I will consistently use quotation marks for such metaphorical evolutionary shorthand) in making the living bodies they grow for their own "use" as perfect and successful as possible. Anything less impairs their own progress toward immortality. One of Dawkins's central points (following Hamilton 1964) is that from a gene's "point of view," a perfect host is one that in a crisis is capable of courting and meeting death to save a sufficient number of close relatives, because they carry that very same gene with known probability. Far from representing a flaw in the life and "design" of that host, such an act proves the host's perfect suitability to and fulfillment of the "purpose" for which that host was grown, namely to reproduce the genes that grew it.

If one switches viewpoint to the more limited perspective of one of these hosts, and explains to them that they had really been "designed" for their glorious act of self-sacrifice on behalf of close relatives by the genes that thereby kept their losses to a minimum in dire circumstances, then that host may feel "used" by the genes that thus furthered their own "interests" at the expense of his or her life. This is how the issue of a self-interest on the part of genes standing in opposition to that of their host arises. It is squarely predicated on the reality of genes, which furnish a causal entity to which a host-transcending self-interest can be attributed, albeit metaphorically. These macromolecular templates (DNA) slip from generation to generation through fertilized ova, growing from them the bodies they ride to the next sexual encounter, down the generations over aeons toward immortality.

In this relation it is the genes that are the enduring entity, causally efficacious in growing the dependent entity, a host body or "disposable soma" (Kirkwood 1977; Kirkwood & Rose 1991). That causal precedence gives them causal and logical priority over the living bodies they grow, and makes it possible to define a "self-interest" on their part relative to the hosts they "bud off" along their way. In the case of goals, on the other hand, there is *no* analogous causal nexus: they perish with the individual who entertains and pursues them. This makes them host-dependent in a way that genes are not, and leaves no causal continuity or mechanism analogous to the transmigrating genes on which to base the conceit of a host-independent self-interest on their part.

With regard to goals, the roles in fact are reversed: here, the host is the more enduring entity, for whom goals come and go, some ephemerally, some cyclically, and some others yet more enduringly, but none with a life expectancy beyond the life of the host. One might in fact analogize, albeit loosely, the succession of goals adopted ("grown") by hosts in furtherance of their prospects during their finite careers on earth to the succession of host bodies the genes grow as instruments in their far longer and potentially endless careers. For goals, it is clearly the striving and erring host that is the causally and logically prior condition, and thus the entity to whom self-interest is to be attributed in relation to goals, and not the other way around, if we are to adhere to the Dawkinsian sense in which the self-interest of his "selfish gene" is defined.

Goals thus lack any causal mechanism such as the transgenerational continuity of genes on which to base host-independence, and with it a causal basis for defining a self-interest apart from that of the host. Should one propose a language-based mechanism of transgenerational transmission of culturally shared goals in this role, one would have to resign oneself to limit the validity of the theory to language-competent humans, and even then it would cover only some of their goals. Excluded would be a vast domain of goals adopted in response to situational happenstance, as well as idiosyncratically adopted personal ones (two sources of evidence cited in the target article), a restriction hardly matching the intentions of H&B. It goes without saying that some of these goals indeed qualify as selfish in the ordinary sense of the word (i.e., pursued at the expense of the self-interest of other *hosts*), but that is not the sense in which "selfish" figures in Dawkins's title, predicated as it is on the host-transcending nature of genes, nor in H&B's attempt to interpret goals in similar terms.

Unless, therefore, we are presented with some other mechanism, as yet unknown, by which the "selfish goal" envisioned by H&B might acquire a host-independent self-interest, it seems safe to conclude that when hosts are "done in" by the goals they pursue, the explanation is not to be sought along Dawkinsian lines, but in more mundane circumstances attending those pursuits. These range from idiosyncratic caprice to mental disorder and include a wide range of ways in which our innate propensities, such as our liking for sugars, can derail when expressed under circumstances different from those under which they evolved. Such matters are topics of active study in behavioral and evolutionary biology and have given birth to the new discipline of evolutionary medicine, where instructive examples can be found (Nesse & Williams 1995; Stearns 2012).

The conscious roots of selfless, unconscious goals

doi:10.1017/S0140525X13002100

Gordon B. Moskowitz[a] and Emily Balcetis[b]

[a]*Department of Psychology, Lehigh University, Bethlehem, PA 18015;*
[b]*Psychology Department, New York University, New York, NY 10003.*
gbm@lehigh.edu
http://psychology.cas2.lehigh.edu/content/gbm4
emilybalcetis@nyu.edu
http://www.psych.nyu.edu/balcetis/

Abstract: We counter Huang & Bargh's (H&B's) metaphoric description of the unconscious, selfish goal on three points. First, we argue, unconscious goals are rooted in conscious choices related to well-being. Second, unconscious goal pursuit occurs through early-stage orienting mechanisms that promote individuals' well-being. Third, unconscious goals work selflessly, resulting in their own demise.

Consciousness is not dispensable. Unconscious goals are not inherently selfish. We suggest unconscious goal pursuit is rooted in conscious choice and occurs through means related to promoting well-being.

Unconscious goals from conscious choices. Unconscious goals are rooted in conscious choices. As Bargh (1990) articulated, unconscious goals originate from conscious learning, deliberation, and evaluation that over one's lifetime can become automatic responses to eliciting situations or cues. Myriad theories support that unconscious goals originate in individuals' conscious decisions. People choose how, under what conditions, and by what means they will satisfy needs. These decisions eventually become routine, automatically and unconsciously activated under similar circumstances (see Custers & Aarts 2010; Kruglanski et al. 2002; Moskowitz et al. 2004; Shah & Gardner 2008). Many unconscious goals are the result of conscious decisions; consciousness is necessary for models of unconscious goal pursuit.

Unconscious means promote the well-being of the individual. Although we agree with H&B that selective attention and biased perception serve as early-stage orienting mechanisms to promote goal pursuit, we take issue with the claim that attention and perception operate in selfish manners. We instead posit that early-stage mechanisms actually facilitate outcomes that generally benefit the individual. For example, people focus visual attention on (Maner et al. 2007), perceive as closer (Balcetis & Dunning 2010), and see as bigger (Van Koningsbruggen et al. 2011) those objects in their environment they are motivated to acquire compared with objects they are not. Similarly, people see objects that pose a threat against which people may need to defend themselves as closer than objects that are either nonaffective or that are disgusting but do not pose a danger that requires immediate response (Cole et al. 2013; Harber et al. 2011). Importantly, unconscious goals systematically direct attention and affect perception in ways that serve individuals' needs (Balcetis & Dunning 2006). Perceiving desired rewards or dangerous threats as focal, close, or large readies the perceiver physiologically and psychologically for goal-related action that ultimately leads to the satisfaction and termination of the active goal state (Balcetis & Dunning 2010; Pichon et al. 2012). For example, people see a bottle of water as bigger and closer, in order to facilitate acquisition of it, and ultimately end their need state of thirst and goal to acquire a drink. Unconscious early-stage orienting mechanisms emerge in ways that promote individuals' well-being, not necessarily the continued activation of the goal.

Selflessness of unconscious goal pursuit. We counter the metaphoric portrayal of unconscious goals as selfish. First, unconscious goals coordinate within a cooperative goal system through multiple processes, including goal shielding. At times, a primary goal inhibits progress on another goal with which it is incompatible. Although dominance of the primary goal may be regarded as selfish, the circumstances under which inhibition occurs suggest it is not. Primary goals do not inhibit secondary goals indiscriminately; instead they often selflessly facilitate compatible secondary goals (e.g., Shah 2003). Inhibition occurs only when secondary goals are antagonistic with individuals' needs and well-being. For example, weight-loss goals inhibit pleasure-eating goals (Fishbach et al. 2003). Egalitarian goals inhibit stereotyping goals (Moskowitz & Li 2011). Studying goals inhibit socializing goals (Aarts et al. 2007). Inhibition is not an act of a primary goal's selfish nature, but an act of shielding an individual from undesired obstacles and distractions.

Second, unconscious goals coordinate by relying on processes that are mutually beneficial for pursuit of multiple goals. Indeed, there are multiple behaviors, choices, or thoughts people can use to attain a single goal – a quality known as equifinality. For example, individuals who hold the goal to be a good kisser might frequently apply lip balm, read magazines for tips, and brush their teeth after every meal. Yet individuals pursue a given goal in manners that can facilitate the simultaneous pursuit of more than one goal. The means selected are those that have compounded value by being able to allow multiple goals to reach fulfillment – a quality referred to as multifinality. Individuals may prioritize tooth brushing to meet the goal to kiss well and maintain dental health. Indeed, goal-directed behaviors are valued highly when they can serve multiple goals and are valued less highly when they serve only a single goal (Chun et al. 2011). Rather than selfishly pursuing the means to one goal at the expense of others, unconscious processes coordinate in pursuit of the simultaneous completion of multiple goals.

To be sure, the metaphor of the selfish goal could be sustained in light of the literature just reviewed. Indeed, selfish motives do drive other forms of prosocial responding (e.g., Cialdini et al. 1987). It is possible that goals hold parasitic relationships with one another. One goal might coordinate with a second goal only to selfishly reap the benefits of the compounded value that such coordination affords.

Nonetheless, a third literature leads to the conclusion that goals are not selfish but instead suicidal. Rather than attempting to preserve themselves, goals seek to end themselves. To be sure, goals are strong; they remain accessible as time passes (Bargh et al. 2001) and do not dissipate if disrupted (e.g., Zeigarnik 1927). However, goal striving decays quickly after individuals attain their goals (e.g., Cesario et al. 2006; Förster et al. 2007; Martin & Tesser 2009; Moskowitz 2002; Moskowitz et al. 2011; Wicklund & Gollwitzer 1982). That is, goals die once completed. They do not seek to selfishly propagate their own existence, but instead seem compelled to work toward self-termination and, in so doing, deliver well-being and need-fulfillment.

Summary. H&B thoroughly review the principles that underlie unconscious goal pursuit that have overhauled psychological characterizations of human motivation. We question, though, whether it is appropriate to cast aside consciousness and to brand unconscious goal pursuit as inherently selfish. Unconscious goals promote well-being through early mechanisms, in coordination with other goals, at the risk of their own demise. Unlike genes, goals are inherently unselfish. They are servants of the higher-order needs of the individual, selflessly working to promote well-being.

Winner takes it all: Addiction as an example for selfish goal dominance

doi:10.1017/S0140525X13002112

Christian P. Müller and Davide Amato
Department of Psychiatry and Psychotherapy, Friedrich-Alexander-University Erlangen-Nuremberg, 91054 Erlangen, Germany.
christian.mueller@uk-erlangen.de **davide.amato@uk-erlangen.de**
http://www.psychiatrie.uk-erlangen.de/wir_ueber_uns/mitarbeiter/prof_dr_rer_nat_christian_p_mueller/index_ger.html

Abstract: Here we argue that the selfish goal concept may well be suitable to explain inconsistencies not only in micro-behaviors, but also in the gross behavioral repertoire of an individual, which is often associated with psychopathologies, such as addiction. Neurophysiological evidence for pathological conditions like addiction emerged, and this evidence may also serve as an explanatory model for normal behaviors.

In their target article, Huang & Bargh (H&B) suggest an interesting comparison between genes and goals. Inspired from the selfish gene theory, which proposes that genes are in competition for survival and reproduction and that the phenotypic organism only serves as a "survival machine" (Dawkins 1976), they inquire in how far a similar framework might work for behavioral goals. The starting point for this argument appears in the observation that humans can pursue different goals. Because we cannot observe "goals" directly, we have to infer them from observing goal-directed behaviors. Individual goals develop and try to get access to behavioral output, which results in a behavioral phenotype. Importantly, there is usually only one behavior that can be emitted at a time, no matter how many goals are established in the organism. It appears all too obvious that there must be some sort of competition between distinct goals for access to behavioral output, in particular because single goals can be counteracting and resulting in behaviors that rule out one another. This goal competition was now taken to a new level by H&B by assigning every goal an innate "selfishness." Thus, goals are selfish and "fight" for their access to behavioral output. H&B argue that this competition is the reason for often inconsistent behaviors with highly disadvantageous components for the individual.

Here, we ask whether the selfish goal concept may well be suitable not only to explain inconsistencies in micro-behaviors, but also in the gross behavioral repertoire of an individual. This is often (but not always) associated with psychopathologies. In a human being with a "normal" behavioral repertoire, one can observe a certain number of goals acting with a balanced influence on behavioral output. This results in a behavioral phenotype that displays a number of distinct goal-directed behaviors. The frequency of each of them depends on whether they are rewarded or punished by the environment. Thereby, an increase in the frequency of one behavior usually means a decrease in the frequency of other behaviors, which may be seen as a behavioral competition. This competition of behaviors may directly reflect a competition of the goals at neuronal level in the brain. Although a majority of humans can maintain a certain amount of rewarded goal-directed behaviors, for some individuals this balance becomes unusually polarized. In that case one "very selfish" goal has taken control over the behavioral repertoire. Likewise, at neuronal level one goal representation might inhibit all others in a "collateral inhibition" for access to behavioral output. This imbalance may occasionally result in a "successful behavioral phenotype," for example, in a professional tennis player, who is playing tennis all the time wishing to win as many tournaments as possible, but suppressing most other behavioral alternatives. Quite often, however, selfish goal dominance may result in psychopathologies, such as drug addiction.

Drug addiction is a syndrome in which the seeking and taking of psychoactive drugs has gained higher priority than virtually all other normally important behaviors. In that, drug addiction is a good example of how mostly unconscious goals, which were developed in instrumental and classical conditioning processes (Robbins et al. 2008; White 1996), may dominate conscious goals of becoming abstinent and spending more time with other activities. For drug addiction, there is now a fairly good understanding of how subconscious goals dominate cognitive ones at neurophysiological level. This may, in turn, inform theories of normal goal competition. Addictive drugs can induce a learning-like neuronal plasticity already after a single administration at synapses of the mesolimbic reward system (Saal et al. 2003; Ungless et al. 2001). During repeated drug administration, classical and instrumental learning processes for goal-directed behaviors are established. Thus, stimulus (S)-response (R) associations are formed, which establish the goal of "mental state change" by means of a drug and further elaborated drug use mechanisms (Müller & Schumann 2011).

However, conscious goal representations may also be formed after only a single drug exposition, by establishing an episodic memory of a desirable drug-induced mental state (Müller 2013). At physiological level, there is an increase in dendritic branching and synaptic spine density in neurons of the reward circuitry, but also in neocortical structures (Robinson & Kolb 2004). Importantly, this drug-induced neuronal plasticity prevents establishment of neuronal plasticity induced by natural stimuli learning (Kolb et al. 2003). Here one may speculate whether this may be one site where drug-goals become "selfish," that is, by suppressing establishment of plasticity and learning directed toward alternative reinforces. Drug addiction requires long-term drug consumption with an escalation of the consumption and a loss of control over this behavior (American Psychiatric Association 1994). The loss of control may be seen as a goal conflict or as an imposition of one goal over the other in which the unconscious goal of "harm avoidance" and the conscious goal of "abstinence" lose influence on the unconscious goal of "drug taking" and the conscious goal of "changing the state of mind" (Müller & Schumann 2011).

Why do these goals lose out? At physiological level, there is evidence that the initially outcome-controlled behavior of drug taking eventually becomes a cue-controlled habit. It is no longer dependent on the actual effects of the drugs, but depends mainly on drug predicting stimuli. This was shown to be mediated by a neuroanatomical loop projection from the nucleus accumbens to the ventral tegmental area and from there to the dorsal striatum (Belin & Everitt 2008; Haber et al. 2000). Habits are controlled by the dorsal striatum and are difficult to suppress by neocortical inhibitory projections (Knowlton et al. 1996). Furthermore, in drug addiction, prefrontal function and behavioral control decline steadily, possibly by a reduction of its serotonergic and dopaminergic innervation (Müller et al. 2010; Pelloux et al. 2012). This reduces prefrontal inhibition of striatal circuits and allows for a more compulsive pursuit of habit-like behaviors. In other words, cortically represented goals are weakened and lose their influence on behavioral output while striatally represented goals become superdominant. Altogether, we suggest that these pathways (Belin et al. 2009) may serve as one physiological base for H&B's conscious and unconscious competition of goals that appear indeed "selfish" when their impact on behavioral output is considered.

Unconscious goals: Specific or unspecific? The potential harm of the goal/gene analogy

doi:10.1017/S0140525X13002124

Bence Nanay[a,b]
[a]*Centre for Philosophical Psychology, University of Antwerp, 2000 Antwerp, Belgium;* [b]*Peterhouse, University of Cambridge, Cambridge CB2 1RD, United Kingdom.*
bn206@cam.ac.uk or **bence.nanay@ua.ac.be**
http://webh01.ua.ac.be/bence.nanay

Abstract: Huang & Bargh's (H&B's) definition of goals is ambiguous between "specific goals" – the end-state of a token action I am about to perform – and "unspecific goals" – the end-state of an action-type (without specifying how this would be achieved). The analogy with selfish genes pushes the authors towards the former interpretation, but the latter would provide a more robust theoretical framework.

The central rhetorical device of the target article is the similarity between selfish genes that often work against the interests of the organism that carries them and our selfish goals that often work against our interests. Although this selfish gene/selfish goal analogy is supposed to be merely illustrative (and the authors presumably do not endorse any strong version of evolutionary epistemology), I will argue that the analogy does more harm than good. More precisely, I will argue that Huang & Bargh's (H&B's) definition of goals as "mental representations of desired end-states" (sect. 2.2, para. 2) is ambiguous, and the selfish gene analogy pushes the authors to resolve this ambiguity in the less promising direction.

The definition of goals as "mental representations of desired end-states" is ambiguous between what could be called "specific goals," where the represented end-state is the end-state of a token action I am about to perform, and "unspecific goals," where it is the represented end-state of an action-type that would satisfy my desire (without specifying how that would be achieved). To use an example, if I am hungry, I can have an unspecific goal of filling my stomach with *something* (which is left unspecified) or a specific goal of eating this particular piece of chocolate cake right in front of me. The ambiguity is not just a feature of the definition H&B give, but they also use examples of these two different kinds of goals interchangeably throughout the article.

The specific goal interpretation of the Selfish Goal Theory makes the analogy with selfish genes very straightforward. Goals in this sense are self-contained entities that strive to be fulfilled, often at the expense of our interests. And this sounds very similar to the conception of genes as self-contained entities that replicate themselves. Leaving the huge debate in evolutionary biology and philosophy of biology aside about just how self-contained entities genes are and whether they should really be taken to replicate *themselves* (but see Godfrey-Smith 2000; Nanay 2002; 2011), what really matters for the purposes of the target article is that in only a smallish portion of the examples do H&B use a specific goal in the sense of being self-contained. In the majority of the examples, the goal is very much unspecific and it is specified by mental states that are not part of the goal itself.

The difference boils down to a difference in what general picture of motivation one endorses. If we accept the self-sufficient specific goal picture, then the only mental state that is needed to motivate us to act is this specific goal. If I have a specific goal of eating this particular chocolate cake, this motivates me to act. But here is a more flexible model that H&B often slide into (and they are right to do so). Two things are needed for motivating us to act: an unspecific goal (of, say, wanting to eat *something*) and a separate representation of something edible in front of us. If this unspecific goal and this "action-oriented representation" are combined, the action is performed.

Just what these "action-oriented representations" are (for example, whether they can be perceptual states) I would like to leave open. They are representations of the objects in the agent's environment that could be used to achieve the unspecific goal. They are not themselves "representations of a desired end-state" – they are representations of means of achieving this desired end-state. And they do not themselves motivate us to act. We are only motivated to act if we have both an unspecific goal and an "action-oriented representation" (see Jeannerod 1997, who calls these "visuomotor representations" and Nanay 2013 who calls them "pragmatic representations").

Which picture of motivation should we choose? H&B themselves seem to be conflicted about this – they seem to go back and forth between these two frameworks, depending on the examples they analyze. But the overall selfish gene analogy pushes them towards the self-sufficient specific goal picture. I argue that this is a mistake and the logic of many of their own examples would demand that they use the more flexible unspecific goal plus action-oriented representation picture.

Everything H&B say can be formulated in this framework: The unspecific goal can be unconscious, and the action-oriented representation can also be (and most often it is) unconscious (Dehaene et al. 1998; Goodale 2011; Jeannerod 1997; Nanay 2013). So we get a more nuanced picture about the relation between the conscious and the unconscious processes that lead to action. To put it very simply, both the unspecific goal and the action-oriented representation can be unconscious. And often both of them are.

Sometimes we have an unspecific goal and this influences the action-oriented representation we form (say, we are hungry and look around in the fridge to see what we can eat). Some other time, the action-oriented representation comes first and this triggers the unspecific goal (say, you walk past a café and see a delicious cake in the window that makes you want to eat). H&B themselves analyze those very interactions in the target article, but they can only be made sense of if there are two mental states that can interact in these two different ways (and not just one), that is, in the more flexible unspecific goal plus action-oriented representation framework. They would be better off using this way of framing their claims instead of the more catchy but ultimately misleading (and from an evolutionary biology/philosophy of biology point of view, somewhat suspicious) analogy with selfish genes.

The selfish goal meets the selfish gene

doi:10.1017/S0140525X13002136

Steven L. Neuberg[a] and Mark Schaller[b]
[a]*Department of Psychology, Arizona State University, Tempe, AZ 85287-1104;*
[b]*Department of Psychology, University of British Columbia, Vancouver, BC V6T 1Z4, Canada.*
steven.neuberg@asu.edu
http://psychology.clas.asu.edu/neuberg
schaller@psych.ubc.ca
http://neuron4.psych.ubc.ca/~schallerlab

Abstract: The connection between selfish genes and selfish goals is not merely metaphorical. Many goals that shape contemporary cognition and behavior are psychological products of evolutionarily fundamental motivational systems and thus are phenotypic manifestations of genes. An evolutionary perspective can add depth and nuance to our understanding of "selfish goals" and their implications for human cognition and behavior.

Huang &Bargh (H&B) draw an analogy between selfish genes and selfish goals. Just as genes "selfishly" build organisms to promote their own replication (and not necessarily in the best interests of the organism within whom these genes reside), goals "selfishly" shape cognition and behavior to promote their own attainment (even though this may retard progress toward other goals that individuals might hold dear). The analogy is both clever and useful, and underscores important insights such as the reconfiguration principle (the idea that an activated goal constrains individual's information processing in predictably goal-centric ways). But the connection between selfish genes and selfish goals is not merely metaphorical. An evolutionary perspective on human motivation implies deep connections between the selfish replication of genes and the selfish impact of goals on human psychology. Key principles of goal-directed cognition can be understood more completely, and their implications predicted more thoroughly, when located within an evolutionary perspective.

The human mind has the capacity to generate a nearly unlimited number of goals. Many of these goals have no obvious implications for, or connection to, genetic reproduction (e.g., the goal of reading, and indeed writing, this commentary). But a great many other goals are linked to specific motivational systems that evolved because, in ancestral populations, they promoted affective, cognitive, and behavioral outcomes that facilitated replication of the genes that built these systems. These fundamental motivational systems – which include motives governing food intake, self-protection, disease-avoidance, social affiliation, mate seeking, mate retention, and child-rearing – reflect a relatively small set of specific fitness-relevant challenges recurrent across our evolutionary history (Aunger & Curtis 2013; Bernard et al. 2005; Kenrick et al. 2010).

These fitness-relevant challenges are qualitatively distinct and require distinct kinds of behavioral responses in order to be successfully met. (As many of us are painfully aware, behavioral strategies that facilitate the initial attraction of a mate may be useless when it comes to maintaining a long-term relationship with that mate.) Indeed, behaviors that facilitate progress toward meeting some fitness-relevant challenges may actually retard progress toward others. The context-contingent activation of goal states plays a vital role in promoting adaptive behavior, and it does so by constraining the perceptual, cognitive, and decision-making processes that govern behavioral responses. It does this not just by facilitating specific goal-consistent perceptual attunements and cognitive biases, but also by inhibiting other attunements and biases that might otherwise energize behavioral progress toward other (less immediately pertinent) goals instead. The reconfiguration principle is rooted in the soil of these evolved motivational systems.

The evolutionary perspective is useful not merely because it provides an ultimate rationale for goal-directed constraints on cognition; it is useful because it provides a deductive framework within which hypotheses can be generated (and empirical discoveries made) about exactly how specific goal states constrain specific aspects of cognition. There is an enormous body of empirical literature that documents numerous ways in which the activation of evolutionary fundamental goal states adaptively constrains a wide range of cognitive phenomena, especially in the realm of social cognition (Neuberg et al. 2013; Neuberg & Schaller, in press).

To illustrate, consider the results of several research projects that have proceeded from observations about the specific ways in which specific categories of people have, in ancestral ecologies, posed threats to individuals' fitness. For example, male members of tribal out-groups historically posed a predatory threat to physical safety. In most contemporary human ecologies, this threat is small, yet the self-protection motivational system that responds to predatory threats remains attuned to perceptual cues connoting out-group status, with consequences for social cognition. For example, when a self-protection goal becomes active (even if by incidental events irrelevant to intergroup interactions, such as watching a frightening movie or being in a darkened room), non-black perceivers are especially likely to have danger-connoting stereotypes of African Americans implicitly activated into working memory (Schaller et al. 2003), to misperceive anger in the objectively neutral facial expressions of black men (Maner et al. 2005), and to identify racially ambiguous angry male faces as black (Miller et al. 2010).

These perceptual and cognitive biases are functionally distinct from the perceptual and cognitive biases that emerge when a person becomes concerned with disease. For example, when a disease-avoidant goal is active, individuals are more visually attentive to disfigured faces (Ackerman et al. 2009). They also become more likely to implicitly associate disease-connoting semantic concepts with categories of people who appear superficially to deviate from a subjectively "normal" appearance – including people who are old, physically disabled, or obese (Schaller & Neuberg 2012). There are additional cognitive consequences of a disease-avoidant goal state, which follow from the fact that, historically, many cultural norms and traditions served as buffers against disease transmission (Fabrega 1997). Consequently, when a disease-avoidance goal is active, individuals are more attracted to conformists, judge norm violations to be more morally wrong, and endorse more conservative (i.e., more tradition-preserving) sociopolitical attitudes (Helzer & Pizarro 2011; Murray & Schaller 2012).

Those last results highlight an important point: active goals (such as disease-avoidance) can constrain aspects of cognition that are transparently pertinent to the goal (e.g., attitudinal aversion to people who appear unhealthy), but they can also constrain aspects of cognition that, at first glance, may not seem so pertinent (e.g., conformist attitudes) – except when located within an evolutionary framework.

It is true that many goals are responsive just to the fleeting incentives of the here and now, and may have only minimal connection to evolutionary fundamental motives of the sorts we identified above. But it is also true that many goals *are* products of these evolved, fundamental motivational systems. These goals are not merely analogous to selfish genes. They are, instead, phenotypic manifestations of motivational systems encoded within and built by selfish genes. By carefully considering the implications of the causal connection between selfish genes and selfish goals, we can more expertly predict the consequences that these goals can have on human cognition and behavior.

Goals reconfigure cognition by modulating predictive processes in the brain

doi:10.1017/S0140525X13002148

Giovanni Pezzulo
Istituto di Scienze e Tecnologie della Cognizione, CNR, 00185 Roma, Italy.
giovanni.pezzulo@istc.cnr.it
http://www.istc.cnr.it/people/giovanni-pezzulo

Abstract: I applaud Huang & Bargh's (H&B's) theory that places goals at the center of cognition, and I discuss two ingredients missing from that theory. First, I argue that the brains of organisms much simpler than those of humans are already configured for goal achievement in situated interactions. Second, I propose a mechanistic view of the "reconfiguration principle" that links the theory with current views in computational neuroscience.

The brain mechanisms supporting the achievement of human-level goals are sophistications of the neuronal architectures used by our evolutionary ancestors for situated interaction. These simple organisms had to select adaptive action rapidly; and for this purpose, a brain design based on a staged perception-cognition-action pipeline was probably too slow. A better design could be a basic (but robust) control-theoretic loop with several operations deploying in parallel and influencing each other; for example, an "affordance competition" architecture that specifies and selects multiple potential action plans in parallel under the biasing influence of current goal and motivation contexts (Cisek & Kalaska 2010).

This embodied view encourages seeing all cognitive processes through the lens of action and goal achievement. Here, the main role of perception is signaling opportunities for achieving valuable goals through action, not providing an objective representation of the external environment (Gibson 1979; Proffitt 2006). Memory, too, is in the service of action and goals: it integrates patterns of past interaction with current perception to provide context for goal selection and achievement (Glenberg 1997; Verschure 2012).

From this perspective, the brains of simple organisms are already well configured for goal achievement in situated interactions. More complex cognitive architectures (including human) might be elaborations of this initial brain design, but each would retain the initial design's essential aspects (Pezzulo & Castelfranchi 2007; 2009). However, this view leaves unexplained how higher animals achieve increasingly more complex goals (e.g., distal, abstract) and with higher behavioral flexibility.

One hypothesis is that increasingly more complex forms of goal-directedness might result from the progressive improvement of predictive abilities, which permitted incorporating future events in decision and action control (Pezzulo 2011). Here, I elaborate on that idea and argue that goals "configure" cognitive processing by controlling the predictive dynamics of the brain.

The notions of *prediction* and *prediction error* are ubiquitous in computational neuroscience theories, including those for perceptual processing (e.g., predictive coding), motivation and reward (e.g., reinforcement learning), and action control (e.g., internal modeling) and its dysfunctions (Frith et al. 2000). Ultimately, the brain could be seen as a prediction machine (Bar 2007; Friston 2010; Pezzulo 2008); still it cannot predict everything.

I propose that *the current goals define the relevant dimensions along which predictions are generated and prediction errors are monitored and corrected*. For example, goals determine which expectations have value and should be fulfilled, which events should be predicted and which errors monitored, how prediction errors should be evaluated (e.g., as good or bad), and what should be learned from them. In sum, goals can bias all cognitive processes by channeling predictions toward goal-relevant events. Below I discuss representative theories assigning goals a role in "(re)configuring" cognition and link them within a common prediction-based framework.

In an early cybernetic model of goal-directedness, the *TOTE*, the discrepancy between desired (goal) and actual (sensed) state triggers a goal-directed action (Miller et al. 1960; Pezzulo et al. 2007). Thus, goals determine what errors are monitored and what actions minimize them; but they do not directly modulate predictive and perceptual processes.

Ideomotor theories recognize those links and argue that actions are controlled by goals, which are coded as distal action effects (Hommel et al. 2001). When a goal is selected, goal-related feature dimensions are "intentionally weighted" and have a stronger impact on perceptual processing and response selection (Memelink & Hommel 2013). "Reconfiguration" happens because goal selection enhances the salience of effect-related (goal-related) sensory events and modulates attention to these events. Similar hypothesis are advanced in theories of selective attention and top-down control (Desimone & Duncan 1995; Miller & Cohen 2001).

Goals have also motivational aspects that influence perception and action. A recent theory emphasizes that perception can combine external (sensory) and interoceptive (drive- and goal-related) signals, so that a mismatch between an internal need (e.g., hunger) and sensory stimuli (e.g., no food) modulates the importance of the visual signals (Montague & King-Casas 2007; Pezzulo 2013), and attention can be deployed to goal-relevant events (Mysore & Knudsen 2011). Motivational factors influence action selection, too, using a mechanism that minimizes reward prediction errors through reinforcement learning (Dayan & Balleine 2002) and during planning (Pezzulo et al. 2013; Solway & Botvinick 2012).

All of those seemingly disconnected ideas can be reconciled within the prediction-based framework of *active inference* (Friston 2005; 2010). In this framework, goals correspond to probabilistic (Bayesian) priors at high hierarchical layers of the brain and are achieved by steering actions that minimize prediction errors between the priors (goals) and the current state. Brain hierarchies encode internal models at multiple timescales that give flexibility and context for goal achievement. Perception compares competing perceptual hypotheses by generating predictions at higher hierarchical levels, comparing the predictions with sensory stimuli, and using the prediction errors for hypothesis selection. Doing so permits reducing uncertainty in the sensorium, which is necessary for accurate action. Performing a complete perceptual processing is too costly; however, it can also be unnecessary as only few perceptual hypotheses correspond to highly valued goals. Thus, during perceptual inference, goals control which aspects of the sensorium should be made less uncertain and bias perception toward the states of affairs that realize the agent goals with higher probability (and as a side effect, they produce an "optimism bias", Sharot et al. 2011). In other words, goals can modulate perception and behavior by regulating attention and affordance selection; in the active inference framework, this corresponds to optimizing the *precision* (inverse variance) of the relevant parts of the sensory and proprioceptive data (Feldman & Friston 2010). From this perspective, perception signals where value (i.e., the opportunity to achieve a goal) is and permits picking up affordances and achieving goals.

A challenge for this framework is explaining increasingly more complex, human-level forms of goal-directedness, often linked to prefrontal function and cognitive control. An initial hypothesis is that cognitive control might essentially replicate the active inference scheme, but in a "covert" (simulated) form (Pezzulo 2012). The covert process allows generating goals that need not be achieved immediately but are retained in prospective memory until there is an opportunity to achieve them. The goals generated through this process might use the aforementioned precision-weighting mechanisms to influence cognitive processing over the long timescales required to achieve distal objectives; at least, unless another "selfish" goal takes control.

ACKNOWLEDGMENT

Research funded by the EU's FP7 under grant agreement no. FP7-ICT-270108 (Goal-Leaders).

Cui bono? Selfish goals need to pay their way

doi:10.1017/S0140525X1300215X

David Spurrett

Philosophy, University of KwaZulu-Natal, Durban, 4041, South Africa.
spurrett@ukzn.ac.za
http://ukzn.academia.edu/DavidSpurrett

Abstract: The target article falls short of explaining the phenomena, including motivational conflict, that it sets out to. The two main reasons for this are: (1) It is unclear in what sense goals are "selfish"; (2) We need an account of *how* selfish goals motivate people. If selfish goals are not in the replication business, then what is in it for them? And if they do not offer people something that they want, how do they ever influence what people do?

The proposal in the target article risks not explaining the phenomena, including motivational conflict, that it claims to. There are two related reasons for this. First, it is not clear enough in what sense goals might be "selfish," or what incentives *they* respond to. Secondly, the proposal lacks an account of *how* selfish goals motivate individual people, including how they compete with other incentives to which people respond.

Goals are apparently selfish in a way "analogous" to Dawkins's (1976) use with reference to genes. Dawkins argued that to understand much of biology, it was necessary to take the perspective of the units of heredity. These are "selfish" in the sense that their only interest is in being replicated. Serving this interest sometimes makes demands that oppose the well-being of the individual carrying them. Dawkins focuses mostly on units of biological inheritance – genes – but also speculates that there may be analogously selfish cultural replicators or "memes." In the case of memes, as with genes, selection processes favour efficient replicators, irrespective of whether this serves the interest of the "host." In both cases it is clear in what sense the replicators are selfish and what is "in it for them." The game of life pays replicators with copies of themselves, and the various individuals in which they occur are means to that end. The interests of these replicators are not *appreciated* by the replicators themselves, but are well defined in evolutionary game theory where payoffs are numbers of descendants (Maynard-Smith 1982).

The target article carefully avoids using the term "meme" and makes almost no reference to replication (except in the context of glossing Dawkins's popularization of gene-centric selection). This suggests that the sense in which goals are selfish *analogously with genes* is not the same as Dawkins' existing meme analogy. But if selfish goals do not have a primary interest in replication, *what are their interests?* Goals, we are told, represent "end-states" and succeed by being *pursued* and/or by being *completed*.

This proposal needs to answer two crucial questions. First, what is "in it" for goals such that pursuit and completion constitute an incentive? Second, how might selfish goals compete for control of an individual?

It is difficult to discern an answer to the first question, or even clues as to what it might be, in the target article. If selfish goals are not replicators and are furthermore in some sense separate from the person they occupy, then they can be rewarded neither with copies of themselves nor in whatever subjective utility the person responds to. (In any event, in the latter case they would not be selfish at all – they would simply be *the person's* preferences.)

The second question is, if anything, even more important. The various (finite) degrees of freedom of any human represent a scarce resource that has alternative uses. That is to say the *problem* of behavior allocation is essentially economic (Shizgal 2012). Some of the processes that implement allocation are peripheral and relatively encapsulated, but most are not. The behavior allocations of people are undoubtedly sensitive to costs and payoffs, even if it is a matter of controversy what specific economic model humans instantiate. In addition, there is mounting evidence that contemplating or selecting both desirable and aversive options in a wide range of modalities (including money, delayed money, risky money, food, drink, pain, looking at attractive faces, and social reputation) is consistently associated with activity proportional to behaviorally inferred desirability in a single brain region (for a recent review, see Levy & Glimcher 2012). Independent of this specific evidence, the motor areas of the brain constitute a final common path for control processes, plausibly requiring any candidate deployment of the agent's capacities to compete on the same terms as the others.

These considerations apply to selfish goals. The options available to a person (including end-states of goals) are often composed of complex mixtures of components (money, food, sex, status, etc.) and in different modalities. Their availability could be immediate or delayed, and they can be subject to risk. The costs of options also vary in magnitude and type (effort, money, pain, delay, etc.) and may themselves be multimodal (e.g., including both monetary cost and delay). Making choices in even an approximately efficient way requires trading off these multimodal options by reference to their net costs and benefits. Arguably, solving that problem is a significant part of what brains are *for*. Unless selfish goals are to be epiphenomenal, therefore, they need to compete along with the already recognized sources of subjective utility (or reward, or reinforcement), and they need to do so somewhere along the recognized pathways for behavioral control.

These factors appear to make almost no impact on the argument in the target article, where there is no mention whatsoever of *utility*, *incentive*, *consumption*, *pleasure*, or *risk*, and merely solitary and passing references to *pain*, *reinforcement*, and *reward*. But selfish goals need to *motivate* the agents that they occupy, and to do so they need to pay their way in some kind of incentive to which those agents are responsive.

Finally, it is worth pointing out that even non-selfish goals can come into conflict and that people motivated purely by their own incentives can exhibit inconsistency over time. The most banal of objectives can be mutually exclusive (resting and working, specializing and generalizing). All that is needed to explain some conflict, that is, is that not all desires (even those of a self) can be satisfied. In addition, the ways we price the costs of various ways of being separated from a reward (by delay, risk, effort, etc.) can themselves be a source of inconsistency. This has been most extensively studied in the case of delayed rewards, where there is considerable evidence that humans discount rewards that will be received later (Ainslie 1992; Kable & Glimcher 2007) in a manner that corresponds more closely to a hyperbolic than exponential function. The mere passage of time, that is, can change the ranking of preferences within a single person.

Automatic goals and conscious regulation in social cognitive affective neuroscience

doi:10.1017/S0140525X13002161

Chandra Sripada,[a,b] John D. Swain,[c] S. Shaun Ho,[a] and James E. Swain[a,d]

[a]*Department of Psychiatry, University of Michigan School of Medicine, Ann Arbor, MI 48109;* [b]*Department of Philosophy, University of Michigan School of Medicine, Ann Arbor, MI 48109;* [c]*Department of Physics, Northeastern University, Boston, MA 02115;* [d]*Department of Psychiatry, Child Study Center, Yale University School of Medicine, New Haven, CT 06520.*

sripada@umich.edu
http://sites.lsa.umich.edu/sripada/
john.swain@cern.ch
http://www.northeastern.edu/physics/people/faculty/john-swain/
hosh@med.umich.edu
http://www.psych.med.umich.edu/profile/?linkid=ho
jamesswa@med.umich.edu
http://www.psych.med.umich.edu/profile/?linkid=jamesswa

Abstract: The Selfish Goal model challenges traditional agentic models that place conscious systems at the helm of motivation. We highlight the need for ongoing supervision and intervention of automatic goals by higher-order conscious systems with examples from social cognitive affective neuroscience. We contend that interplay between automatic and supervisory systems is required for adaptive human behavior.

Huang & Bargh (H&B) challenge traditional agentic models that place conscious systems at the helm of motivation. Their Selfish Goal Theory sees an agent as a decentralized population of *goals* – mental representations of desired end-states that are automatically cued by environmental contexts. Importantly, on this picture, most motivated behavior arises from complex interactions between multitudes of goals that are constantly being activated and executed according to situation and context. Conscious systems by contrast are claimed to play a limited and peripheral role in our agentic lives.

We propose that automatic goals and conscious systems are more closely interlinked than Selfish Goal Theory portrays. Unconscious goals may very well be the proximate driver of behavior in most situations. But, in ways that we elaborate below, maladaptive goal pursuit is often avoided because of ongoing monitoring by higher-order systems and availability of conscious intervention for regulation of inappropriate motives.

The relationship between automatic goals and conscious supervisory control is vividly illustrated in cases of temptation. H&B discuss cases in which a person's goals are in conflict. For example, a person has the goal of losing weight, as well as the goal of eating an appealing dessert. More extreme instances are found in addiction, where a person has the long-standing and central goal of never again consuming the drug, as well as the suddenly active goal of getting high. H&B note that goal conflict of this kind can lead to regulation that occurs unconsciously, i.e., automatically and without conscious awareness. Although unconscious or other forms of automatic regulation are certainly real phenomena (Gross 2002; Phan & Sripada 2013), a common feature of temptation-involving situations are hardly discussed by H&B: Conscious, volitional self-regulation strategies are mobilized to inhibit the temptation-directed motives. More broadly, effortful top-down regulation of prepotent, automatic, or situation-cued motives is a ubiquitous phenomenon, and Selfish Goal Theory ought to make a place for it.

We propose a dual picture of motivation that recognizes dynamic interplay between automatic goal structures and supervisory conscious systems (Chaiken & Trope 1999; Hofmann et al. 2009; Sripada 2014). Monitoring systems play a critical role in adaptively linking the two systems. An example is the conflict detection system implemented in the anterior cingulate cortex (Botvinick et al. 2004). This brain regions fires in the presence of discrepant prepotent responses (Barch et al. 2001), performance errors (Carter et al. 1998), or competing goals or judgments (Greene et al. 2004; Seymour & McClure 2008). Activation of this region is thought to bring online supervisory systems, typically conscious, that bring additional cognitive resources to bear in conflict resolution (Botvinick et al. 2001).

Indeed, conscious supervisory systems are implicated in a range of regulatory control processes (Gross 2002). For example, roughly 50 neuroimaging studies have examined the phenomenon of volitional emotion regulation (Ochsner & Gross 2005; Ochsner et al. 2012; Phan & Sripada 2013; Swain 2011). In these studies, participants view or hear aversive stimuli while undertaking regulation strategies such as reappraisal and distancing in order to inhibit spontaneous negative responses. These studies confirm circuits linking prefrontal and superior parietal regions, responsible for planning and higher-order thought, with amygdala and striatum, which are responsible for producing spontaneous aversive and appetitive responses (Phan & Sripada 2013). In another specific example, new mothers exercise regulation in responding to the aversive stimulus of baby-cry, which is correlated with the behavioral construct, parental sensitivity construct (Kim et al. 2011), which in turn – based on key parental brain circuits and hormones that are activated by baby stimuli – facilitates the development of key social brain systems in the child (Mayes et al. 2005; Swain et al. 2004; 2011; Feldman et al. 2013).

Although H&B emphasize the automaticity of goals and unconscious regulation, we believe that conscious, volitional regulation processes such as these emotion regulation processes contribute importantly to resolution of goal competition, especially in the context of temptation-involving situations such as dieting and addiction (sect. 4., paras. 1 and 2).

These observations suggests an alternative picture in which automatic goal systems and conscious supervisory systems might be more interlinked more tightly than H&B suggest. Selfish Goal Theory might well be right that the preponderance of behavior emerges from automatic goal processes. But these processes are themselves at all times under constant monitoring by higher-order systems. In key contexts – goals are in serious conflict; goals of substantial significance are being undertaken; goals are pursued in contexts where error could lead to significant negative consequences – conscious supervisory systems are brought online for additional in-depth evaluation of priorities and for real-time adjustment and control in order to fix errors and increase survival potential. The coordinated operation of both systems is required for an agent to produce adaptive behavior.

ACKNOWLEDGMENTS

Chandra Sripada is partially supported by NIH grant AA020297, Center for Computational Medicine Pilot Grant, and the John Templeton Foundation. John D. Swain is partially supported by the National Science Foundation. James E. Swain is partially supported by an Independent Investigator Award from the National Alliance for Research on Schizophrenia and Depression (Brain and Behavior Foundation), the Michigan Institute for Clinical Health Research and the National Center for Advancing Translational Sciences (UL1TR000433).

Goals are not selfish

doi:10.1017/S0140525X13002173

William von Hippel[a] and Frank A. von Hippel[b]
[a]*School of Psychology, University of Queensland, St Lucia QLD 4072, Australia;* [b]*Department of Biological Sciences, University of Alaska Anchorage, Anchorage, AK 99508.*
billvh@psy.uq.edu.au
http://www.psy.uq.edu.au/directory/index.html?id=1159
favonhippel@uaa.alaska.edu
http://www.uaa.alaska.edu/enri/people/Fellows/vonHippel/

Abstract: The metaphor of selfish goals is misguided. Organisms can be considered vessels that further the interests of their genes, but not vessels that further the interests of their goals. Although goals can act at cross-purposes to each other and to longevity, such trade-offs are predicted by evolutionary theory. The metaphor of selfish goals provides no purchase on this problem.

Huang & Bargh (H&B) argue that goals are "selfish" in the sense that Dawkins (1976) proposed that genes are selfish. Although organisms can reasonably be considered vessels that further the interests of their genes, they cannot be considered vessels that further the interests of their goals. Goals serve the organism – not vice versa – and goals have no life beyond the organism (i.e., they have no informational component that is transmitted across generations). Organisms evolved to have goals with survival or reproductive value because such goals improve the organism's fitness, but unlike genes, goals cannot serve themselves.

The problem with the selfish goal metaphor becomes apparent when H&B use it to guide their interpretation of research on goals. For example, H&B write, "the tension between the behavioral imperatives issued by the currently active goal and the other priorities of the person pursuing that goal (over time and across situations) can produce trade-offs between what is 'good' for the goal being pursued versus what is 'good' for the individual" (sect. 4, para. 2). Such statements do not bring us any closer to understanding goals and their interactions because nothing is "good for a goal." Rather, some goals are achieved and others are not, and achievement of some goals (e.g., satisfying a desire for sweets) will lead to failure of others (e.g., satisfying a desire to lose weight).

This problem of how to interpret mutually incompatible goals is compounded by H&B's tacit assumption that an individual's paramount goal is to enhance survival. For example, H&B note that young women with a mating goal are more likely to enhance their attractiveness through diet pills and tanning booths, despite the fact that such activities "operate to the long-term detriment of the individual" (sect. 4, para. 3). H&B interpret the fact that mating goals can undermine longevity as evidence that such goals are acting selfishly.

The confusion here lies in H&B's assumption that survival goals act in the best interest of the individual and mating goals do not. From an evolutionary perspective, however, survival is only important to the degree that it facilitates reproduction – reproductive success, not survival, is the currency of natural selection. Indeed, this is a clear implication of Dawkins's selfish gene argument (and later arguments about selfish genetic elements; Burt & Trivers 2006). Thus, both survival and reproduction are in the interest of the individual, but because the factors that enhance one are not always the same as those that enhance the other, there will be times when the two goals are in tension. When reproduction goals can be enacted well before there is any cost to survival goals – as in the case of tanning and excessive dieting – the individual can be counted on to sacrifice long-term survival in pursuit of reproductive opportunities.

This effect can be seen in a wide variety of biological trade-offs, including senescence itself. Organisms age and die in part because selection favors the allocation of resources to reproduction over tissue maintenance and repair. Because selection pressures diminish with age (Medawar 1952), traits that are deleterious for long-term survival can flourish if they have a positive effect on reproduction. An intriguing example of such an effect can be found in the ε4 allele of the ApoE gene. This allele leads to a greater likelihood of developing Alzheimer's disease late in life (Corder et al. 1993) but, ironically, is associated with better attention and memory early in life (Han et al. 2007). As such examples demonstrate, the goal of maximizing reproduction typically wins out over the goal of maximizing longevity, thereby causing genetic variants such as ApoE-ε4 to be relatively common.

Indeed, because reproduction goals are paramount, the individual will often risk immediate survival in pursuit of reproductive opportunities. For example, men take greater physical risks when in the presence of attractive women, and this rise in risk-taking is mediated by increases in testosterone (Ronay & von Hippel 2010). Similar effects can be seen in many other animals. Because men have much greater reproductive variability than women, and because they are much more likely than women to leave behind no offspring at all, selection favors men who will risk their survival in service of the goal to reproduce. Such findings are not evidence that goals are acting selfishly at cross-purposes to the individual who holds the goal. Rather, they are evidence for sexual selection; genes that cause the organism to behave in a manner that facilitates reproduction are more likely to increase in frequency over generations. This includes genes that result in more offspring at the expense of longevity.

If goals are not selfish, then how are we to understand the goal conflict that H&B highlight? We suggest that the research reviewed by H&B is better understood as individuals choosing between competing opportunities than as the goals competing themselves. Perhaps the most common source of goal conflict occurs when short-term and long-term goals collide. When people choose between short-term gains and long-term costs, they often seem to be acting against their own self-interest in a manner that suggests selfishness on the part of short-term goals. But economists and evolutionary theorists agree that future opportunities must be discounted because of their inherent uncertainty, and thus there is strong and recurrent selection pressure to weight the present more heavily than the future. Such selection pressures are misapplied in cases like drug use, when the good feelings that are produced by drugs hijack a motivational system that evolved to induce organisms to pursue their basic needs (e.g., for social status; Morgan et al. 2002). Nevertheless, none of these effects require that we hypothesize that the goals themselves are selfish, and indeed such an analogy is more likely to lead researchers astray than to serve as a useful metaphor for evolutionary research on goal pursuit.

Should an individual composed of selfish goals be held responsible for her actions?

doi:10.1017/S0140525X13002185

Natalia Washington and Daniel Kelly

Department of Philosophy, Purdue University, West Lafayette, IN 47906-2098.
nataliatwashington@gmail.com **drkelly@purdue.edu**
http://web.ics.purdue.edu/~drkelly/

Abstract: We discuss the implications of the Selfish Goal model for moral responsibility, arguing it suggests a form of skepticism we call the "locus problem." In denying that individuals contain any genuine psychological core of information processing, the Selfish Goal model denies the kind of locus of control intuitively presupposed by ascriptions of responsibility. We briefly consider ways the problem might be overcome.

Recent work has sparked various forms of the worry that the facts of human cognitive makeup lead to skepticism about free will, moral responsibility, and whether reflective deliberation and conscious decisions actually influence behavior (Doris, in press; Levy 2012; Nahmias 2010; Roskies 2006; Saul 2013). Rather than contest the substance or details of the target article, we assume for discussion that it is on the right track and draw out one implication for common practices of holding ourselves and each other responsible for our behaviors. We argue that the picture of an individual person as fragmented into a cluster of autonomously operating selfish goals competing for psychological resources suggests an interesting kind of problem for responsibility ascriptions.

Huang & Bargh (H&B) marshal an impressive amount of empirical evidence in support of their Selfish Goal model, which they articulate in terms of four principles (autonomy, reconfiguration, similarity, and inconsistency). For our purposes, two features are most important, and both focus on control and behavior. First, although the model depicts an individual person as psychologically fragmented (rather than integrated or unified), it does not suggest that the individual's actions will be utterly random or under no control whatsoever. Nor does it suggest that those actions will be completely detached from internal psychological processes. Rather, a specific episode of behavior will ultimately be under the control of a specific goal, namely, whatever goal has become active and won the competition for access to the lower-level machinery closer to the behavioral periphery. Importantly, goals, in H&B's view, are where the buck stops; they are the highest level, most sophisticated psychological states that enter into the production of overt behavior. There is no higher court to appeal to, no further, more global mental entity that selects among the goals, unifies the individual psychologically, or serves as a more constant source of control over different goal-driven episodes of behavior (hence, the propensity for inconsistency that they stress). Second, although H&B do not deny the existence of a self, they allow it only a severely limited function, relegating it to the politician or public relations role of constructing rationalizations for actions that the self has no hand in selecting, producing, or controlling. With the exception of linguistic behavior, the self is behaviorally epiphenomenal.

This picture appears to be at odds with much everyday thought, including the thinking about and practices surrounding moral responsibility. According to common practices, an individual is responsible only for those behaviors over which she has proper control (as opposed to behaviors that are coerced, accidental, or brutishly reflex-like). When she successfully wields such control, the resulting behavior is an appropriate target for responsibility ascriptions. Although there is not yet any clear consensus on the exact form of control required, a recognizable theme is that individuals contain a stable, continuous psychological core that is the source of this control when it is properly exercised. That is, control is a two-part relation, with one relata being the behavior, and the other what we will call the individual's psychological *locus*. With the term "locus," we are trying to remain neutral on whether one thinks of this psychological feature in terms of an individual's self, character, identity, Cartesian mental substance, immortal soul, or whatever. However the notion of a locus is cashed out, typical responsibility ascriptions presuppose that individuals have this kind of central psychological feature. In other words, in ascribing responsibility for an action, there is something we are ascribing responsibility *to* – some central element of the individual that is the ultimate source of the behavior and to which responsibility attaches.

This then is the worry: the Selfish Goal Theory suggests that individuals contain no such psychological locus, because our behaviors are ultimately determined by a loose collection of autonomously operating goals, each with its own agenda. From the perspective of moral responsibility, neither praise nor blame for behavior will be justified because psychologically there's no there there.

We do not think the locus problem is insurmountable. Even if the Selfish Goal model is correct in denying the existence of a traditional psychological locus, there are reasons to think individuals do have the tools to act as responsible agents. An increasingly prominent movement in philosophy of mind emphasizes the embodied, distributed, social, and externalized character of much cognition. Thinkers like Dennett (2003) and Clark (2008) stress that even sophisticated human behaviors can be guided by decentralized control systems that lack a stable or continuous core (see also, Shapiro 2007). Two recent elaborations on these themes deal more directly with moral responsibility and agency.

Doris's (in press) *dialogic conception of agency* recasts the post hoc and socially mediated rationalizations that are the domain of H&B's "conscious self" as in part *constitutive* of human agency. For an individual to participate in this kind of moral reasoning allows that individual's behaviors to become self-directed, or guided by those goals relevant to her most important values. Thus, an individual can justifiably be held responsible for a behavior even if that behavior is not under the control of, say, her internal locus of reflective deliberation.

Ismael (2007; 2011) argues that selves should be thought of in terms of *self-governing systems.* For her, the conscious self is not merely epiphenomenal, but part of a feedback loop that evolved to help organize and guide sophisticated behaviors. Here, the self is not a centralized controller; rather, it is a distributed but higher-level subsystem of the mind "perfectly compatible with a fully decentralized understanding of lower-level processing" (Ismael 2007, p. 87).

Both projects point to conceptions of agency and behavior control that do not rely on the existence of a traditional, stable, centralized psychological core. If there is such thing as a "real self" on these views, it is a dynamic, distributed sort of thing. Although it does not guarantee complete consistency in our behavior, it may get us what responsibility ascriptions seem to require. By our lights, both look well suited to accommodate both our everyday practices of holding individuals responsible and the Selfish Goal's denial of a psychological locus.

Authors' Response

doi:10.1017/S0140525X1300352X

Multitudes of perspectives: Integrating the Selfish Goal model with views on scientific metaphors, goal systems, and society

Julie Y. Huang[a,b] and John A. Bargh[c]
[a]*Rotman School of Management, University of Toronto, Toronto, ON M5S 3E6, Canada;* [b]*College of Business, State University of New York at Stony Brook, Stony Brook, NY 11794;* [c]*Department of Psychology, Yale University, New Haven, CT 06520.*
julie.huang@rotman.utoronto.ca
john.bargh@yale.edu
www.yale.edu/acmelab

Abstract: In our response, we address commentators' feedback regarding the contributions and limitations of the Selfish Goal model. We first clarify potential misunderstandings regarding the model's contributions and the role of consciousness. Second, we situate evaluations of the selfish metaphor within the similarities and differences inherent to the goal-gene comparison. We then respond to commentators' insights regarding future directions and implications of our model, particularly with respect to the broader organizational systems in which goals may operate. Finally, we reiterate important considerations for goal research moving forward.

If we presented one voice in the target article to describe a goal-based model of behavior, then it seems the article elicited (following Whitman) multitudes of perspectives. Taken together, the twenty-six expert commentaries are a mirror that reflects back to us important complexities and debates inherent to the study of goals. We thank all the commentators for their expert feedback; we have learned tremendously from them.

In our response, we emphasize how the commentaries' diverse viewpoints complement each other, and the target article, more than they conflict. We clarify what the Selfish Goal model addresses (common goal structure) and what it does not evaluate (specific functions of consciousness). Critiques of the "selfish" metaphor are addressed by describing how both the similarities and the incongruities offered by the goal-to-gene analogy have the potential to inform future goal research.

With help from the commentaries, we then situate the model's contributions within motivational science and elaborate on additional mechanisms that may drive selfish goal behavior. We also respond to the theme of goal dynamics, which emerged in the commentaries, and make the preliminary observation that control systems theory may prove particularly instrumental when incorporating lateral and hierarchical goal relationships into our model of selfish goals. To conclude our response, we discuss the implications of the present framework as it applies to individuals within social institutions and environments.

R1. Clarifications about the role of conscious and unconscious processes

In the target article, we argue that goals are mental representations that operate in ways that can be described as selfish; consequently, the exact functions of consciousness fall outside this argument. Some commentators (e.g., **Moskowitz & Balcetis**) appear to have misinterpreted the article as making the claim that conscious processes are "dispensable"; to the contrary, we proposed a model that assumes a key role for the integrative functions of conscious states (sect. 2.1).

Part of the misunderstanding may have arisen from our closing remarks regarding the key role of conscious thought in creating rationalizations and managing inconsistencies (sect. 5.2), and we take this opportunity to clarify our intent. We wanted to (1) refer readers to one prominent idea about the function of consciousness (i.e., Gazzaniga 2005) and (2) highlight that conscious understanding of unconsciously rooted behaviors can be mistaken, with the behavior interpreted in terms of plausibility rather than insight (our use of the word "rationalization" emphasized that accuracy is not a guaranteed part of this process; see Bar-Anan et al. 2010).

It is also possible that misreadings of our intent may reflect the influence of another related and topical research tradition that questions the extent to which certain psychological phenomena are conscious or unconscious (e.g., Newell & Shanks 2014). However, we highlighted evidence for the existence of unconscious processes (sect. 3.2) and similarities between conscious and unconscious goals (sect. 3.4) to underscore a different argument – namely, that many interesting and theory-relevant similarities exist between the two forms of pursuit, and this knowledge should be brought to bear on current views in motivational science. With this clarification, we hope to minimize potential future misreadings of our model and refocus the conversation toward the structure and interaction of goals (both conscious and unconscious) as they organize a person's behavior.

R2. Does selfish = useful?

We especially appreciate the opportunity to discuss the "selfish" metaphor we borrowed from Dawkins (1976) to describe the relationship of goals to the person pursuing them. The analogy struck a chord with reviewers, whose responses ranged from enthusiastic support to strong reservations. At the early stages of model development, we believe a metaphor's usefulness in behavioral science might be judged by how it calls attention to key similarities and differences that guide further theorizing and study, and casts fresh eyes within an area of research (Gentner & Grudin 1985; Hoffman et al. 1990). The analogy employed in the target article helps accomplish both tasks, and it also addresses other concerns voiced by the commentaries.

Specifically, the selfish metaphor conveys a number of key similarities between goals and genes, including notions of independence, the existence of a smaller "selectable" unit of analysis, multiple autonomous units co-existing within a single entity, and competition and cooperation between units. Many commentaries responded positively to these concepts and the model inspired by the analogy (e.g., **Becker & Kenrick**; **Conroy-Beam & Buss**; **Cuzen, Fineberg, & Stein** [**Cuzen et al.**]; **Eitam & Higgins**; **Hirsh**; **Kay & Jost**; **Mazzone**; and **Neuberg & Schaller**). Their insights, and the promising directions they offer for future research, are discussed in sections below.

Interestingly, clinically oriented commentators were especially likely to endorse the analogy and/or the model it inspired. This may be – at least partially – because multiple clinical interventions employ a similar focus on the mechanistic underpinnings of behavior when trying to effect change in the patient him or herself. For example, the disease model of addiction de-emphasizes the extent to which patients' destructive behavioral patterns (e.g., gambling) are framed as issues of personal responsibility or shame. Therapeutic interventions highlight the specific steps that a patient must follow to fix his or her current state (regardless of how that patient got there). Other clinical approaches, such as motivational interviewing, similarly shift focus away from substance use as a fundamental issue of willpower (or lack thereof), and instead encourage patients to understand the costs and benefits associated with using, and those involved in quitting. Consequently, our model may have especially appealed to researchers who regularly develop interventions based on the nuts and bolts of behavior.

It is less clear to us, however, why the utility of the selfish metaphor should hinge on questions of evolutionary biology. **Mattei** criticized our model for relying on the empirical validity of Dawkins's original selfish gene theory, which is "at best, [an] unproven biological paradigm." In our target article we specified that the comparison between goals and genes was an analogy and that questions regarding specific mechanisms of evolution fall outside the scope of our argument.

A number of commentators voiced more serious concerns that the metaphor was misleading because discrepancies exist between a gene's interests (vis-à-vis the individual) and a goal's interests (e.g., **Fedyk & Kushnir**; and **Merker**). For example, Merker noted that goals cannot be selfish exactly as genes are said to be selfish, because genetic self-interest is defined by its host-independence (ability to exist independent of the individual organism), whereas a goal cannot exist outside of the person who pursues it. **Spurrett** additionally questioned how goals are to be "rewarded" as genes are rewarded through replication along a similar vein.

It is important to note that even imperfect analogies can prove useful if the dissimilarities help generate new ideas for research (e.g., Cornelissen 2004; Hoffman et al. 1990). Social science provides examples where both similarities and dissimilarities between compared domains lead to productive lines of research. For example, natural selection within population ecology has provided a useful platform to conceptualize organizational behavior (Amburgey & Rao 1996; Hannan & Freeman 1977). Understanding that organizations succeed or fail *as if* they were living beings highlights key similarities between the domains such as dependence on the environment, competition for limited resources, and stages of existence which relate to health and age. Of course, organizations are not exactly like species; species adapt through genetic mechanisms, whereas organizations change through learning processes. Knowing the mismatch exists, however, inspired the eventual identification of "active" and "passive" learning processes unique to organizational adaptation (e.g., Stopford 2003).

The strength model of self-control (Baumeister et al. 1998) also provides examples in which evoked similarities and dissimilarities serve as foundations for productive research lines on willpower (Baumeister et al. 2007; Hagger et al. 2010). Conceptualizing self-control as a muscle nicely captures similarities between the two domains, including dependence on a limited resource, possibilities for short-term impairment, vulnerability to extreme usage, and potential improvement given long-term training. Moreover, the dissimilarities between willpower and muscle-power are also interesting. Biological resources do not explain self-regulatory abilities exactly as they do physical acts, leading Job and colleagues (2010; see also Molden et al. 2012) to discover that participants' lay theories about willpower have the unique potential to predict whether they will be successful at self-regulatory tasks. Importantly, the authors themselves saw these results as offering promising lines of future research, rather than necessarily challenging the overall utility of the muscle metaphor.

Just as organizations adapt in a different manner than species-related change, and willpower is distinct from muscle strength, a goal's selfishness will differ from the exact definition of self-interest as applied to genes (Dawkins 1976). Consequently, although we acknowledge **Merker**'s point that a goal cannot be self-interested exactly as Dawkins defined it within the gene-individual relationship, we nevertheless maintain that an appropriately contextualized concept of goal-selfishness provides directions for insights into motivational science.

Specifically, key comparisons should be made between genes and the (physical) individual organism, and goals and an individual's corpus of behavior. Appreciating that any selfishness of goals must occur within an individual's corpus of behavior (as opposed to across generations of individual organisms) focuses our inquiry on observable replication-like effects within a person's lifetime. Indeed, research suggests that a long-term consequence of achieving a goal is that one typically becomes more strongly motivated to pursue that goal in the future (as held by self-efficacy and other theories; e.g., Bandura 1977). A combination of effects associated with success at goal pursuit can theoretically produce a goal-perpetuation effect, or literally a replication of the goal into the future, in which goals that are successful become stronger, whereas those not as successful became weaker based on performance feedback from the environments frequented by the person. Put another way, what is "in it" for goals is to be pursued more often; and the prediction that success at a given goal pursuit begets future successful iterations is in harmony with other motivational models. As held by self-efficacy theory and supported by more recent research using affective conditioning by Custers and Aarts (2005), the positive affect associated with goal attainment increases the incentive value of that goal in the future for the individual. Meanwhile, the subsequent formation of situational cue-goal associative linkages simultaneously increases the relative likelihood of pursuit (e.g., Bargh 1990; Gollwitzer 1999; Veltkamp et al. 2008) and enables readiness in precisely those situations in which the individual has historically encountered success at pursuit.

More generally, the selfish metaphor provides a newer perspective on current goal research, refocusing attention from possibly unresolvable debates regarding the relative strength of one goal compared to another (**Kay & Jost**). The analogy also links areas of research that are less easily incorporated into commentators' alternative models. Whereas **Moskowitz & Balcetis** suggested that unconscious goals are all originally conscious and become unconscious only from extensive experience (to then ultimately promote an individual's well-being), that alternative runs counter to the considerable evidence and theory regarding innate, evolutionarily acquired adaptive goals (**Becker & Kenrick**; **Conroy-Beam & Buss**; and **Neuberg & Schaller**), infant motivations acquired prior to any experience (e.g., Baillargeon et al. 2013), and extremely negative outcomes of addiction-related disorders (e.g., **Cuzen et al.**; **Müller & Amato**). Moreover, interpreting the goal literature from this "conscious first" perspective, or one which sees "individuals choosing between competing opportunities" (as suggested by **von Hippel & von Hippel**) faces the additional hurdle of explaining how a person chooses and actively pursues opportunities that promote well-being when he or she is unaware of having, or actively pursuing, the goal.

In contrast, conceiving of goal influences as operating selfishly offers a reliable way to describe how a person instigates, guides, and stops processes of which he or she might not be aware (because these processes do not require individual awareness for their operation). Moreover, a goal-driven perspective helps account for research that suggests even when people consciously instigate goals, they are not necessarily in control of all operational features of pursuit, such as which environmental stimuli will be affected by pursuit and when the end-state is attained (Bargh et al. 2008).

Animate-being metaphors (where "ideas or aspects of the mind are likened to creatures," as when goals are described as selfish; Gentner & Grudin 1985, p.184) were frequently employed in psychological discourse during the turn of the century, but they disappeared with increased popularity of other types of metaphor (e.g., computer systems metaphors; Gentner & Grudin 1985). We concede that the relative novelty of an animate-being metaphor highlights the need for interpretive caution, but we are far less convinced of **Mattei**'s argument that it represents a return to the "naïve intellectual modus operandus" that earlier plagued prescientific reasoning. Although we are indeed in early stages, the merits of our model should be judged according to how it helps conceptualize findings in recent research and points to areas that need further theorizing and study.

R2.1. Goals as critical pieces of the whole

Von Hippel & von Hippel (with similar points raised by **Mattei** and **Merker**) questioned the theoretical novelty of the Selfish Goal model, observing redundancies between examples of selfish goal behavior and evolutionary trade-offs. Those commentaries correctly point out that some research cited in support of our model (e.g., studies that find an active mating goal changes how an individual prioritizes physical self-protection) are already understood as cases of sexual selection favoring suboptimal individual-level outcomes.

If the target article had sought to inform evolutionary theory, that critique would be fair; however, it did not. The selfish goal framework is less concerned with isolating when goal conflicts are likely to result in evolutionary trade-offs, or identifying which goals are ultimately best for the individual in those conflicts, than it is with examining the *structural* commonalities that span the pursuit of all goals (including, but not limited to, those involved in sexual selection). Toward those ends, evolutionarily adaptive goals simply provide powerful demonstrations of the reconfiguration principle. We cited experimental demonstrations of mating imperatives conflicting with self-protection goals to illustrate how people's attitudes can shift based on the temporary goal they are pursuing – even in cases where pursuit of that goal results in arguably negative health consequences.

Consequently, we agree with **Merker** that our target article concerns the "more mundane circumstances attending those [goal] pursuits" – although we might have used the adjective "diverse" instead to emphasize the importance of context in the study of goal structure. Charting how goal operation unfolds across multiple, everyday circumstances helps reveal fundamental patterns to its influence over individual human perception and behavior. To take as an example: the self-protection goal is associated with functional changes to perception and behavior that, in ancestral environments, increased the likelihood that a person would avoid potential threats (e.g., being quicker to notice and sidestep a snake in one's path; Kenrick et al. 2010; Neuberg et al. 2004). In current-day circumstances, people react in similar goal-functional ways toward stimuli that are far removed from actual harm (e.g., being quicker to pull a joystick away as the word "snake" is flashed across a computer screen). The wide breadth of goal pursuit, as it spans contexts of varying ecological validity, indicates the presence of general operational patterns. The Selfish Goal model predicts similar outcomes for evolved goals in past and current contexts, as well as for nonevolved goals in the current context – because, as noted by **Neuberg & Schaller**, goals today operate through the reconfiguration principle, just as they did in the evolutionary past.

Goal conflict is mentioned often in the target article, which may have led **Ainslie** to conclude that examples of selfish goals conflicting (e.g., eating a cake versus exercising) are already understood under the framework of intertemporal choice theory. It may be the case that many of our examples involve temporal conflict, such that short-term benefits are weighed more heavily as compared to long-term benefits; however, in everyday life, a person may feel conflict over goals of equivalent temporal qualities – as when one chooses between different hedonic experiences, different future courses of action, or goods associated with different motivations. Moreover, the Selfish Goal model offers additional predictions regarding how the chooser will see the world and act before the choice (becoming quicker to perceive and physically approach stimuli related to the most incentivized goal regardless of when specific benefits associated with goal pursuit can be realized) and immediately afterward (inhibition of mental constructs associated with pursuit).

R2.2. Pathways to selfishness

We are especially grateful to the commentators who were able to anticipate others' concerns regarding the limited treatment of neurophysiological mechanisms in our target article. **Bliss-Moreau & Williams**, **Cuzen et al.**, **Müller & Amato**, and **Pezzulo** elaborate on the affective, neuropsychological, and computational mechanisms that contribute to the selfish pattern of pursuit. In so doing, they provide welcome contributions to the open task of incorporating the Selfish Goal model "along the recognized pathways for behavioural control" (**Spurrett**).

In the target article, we described a sequence of events whereby the more frequently a goal is pursued, the more likely it is to be incentivized for future pursuit (sect. 3.2.1); **Bliss-Moreau & Williams** nicely expanded on this topic. Their commentary addresses how positive affect (incentive) becomes associated with the end-state of that goal, which consequently deprioritizes other goals within the person's repertoire, in accordance with views in many motivational models that affect serves key functions across specific stages of pursuit including goal formation, phenomenal experiences of failure and success (e.g., Kruglanski et al. 2002), and goal switching (Carver & Scheier 1998).

Müller & Amato home in on the striatal circuits and physiological processes in which the consequences of goal conflict can be observed. In extreme cases such as with addicts, drug consumption increases learning-like neuronal plasticity (e.g., increased dendritic branching in the brain's reward systems) for that pursuit. The mechanisms that typically reinforce the pursuit of other goals (e.g., establishment of plasticity and learning in the reward circuits) are prevented, thus leaving the more incentivized addiction goal to dominate. We concur with Müller & Amato's speculation that this chemical process may occur to a lesser degree even for the goals that are not as "unusually polarized" as those in substance addiction, and provide an example of "selfishness." Some selfish goals may be strong enough to block physiological mechanisms that typically reinforce the operation of other goals.

Cuzen et al. forge a similar connection to the literature on impulsive and compulsive disorders. They observe that predictions derived from the Selfish Goal model are very similar to the outcomes observed in clinical populations when the habit system dominates behavior. Individuals with obsessive-compulsive disorder (OCD) bear out the inconsistency principle by acknowledging that their compulsions are unreasonable or harmful even as they continue engaging in those behaviors. Repetitive behaviors can be precipitated by exposure to an environmental cue and, once triggered, are difficult to control, consistent with the automaticity and reconfiguration principles.

Pezzulo translates the reconfiguration principle into the language of computational neuroscience. Specifically, during a phenomenon known as active interference, priors (or goals) have a similar "selfish" influence upon an individual by directing which potential future events have value (as in Ferguson 2008), which prediction errors should be monitored and evaluated, and the lessons that should be taken from events related to successful and unsuccessful prediction of the future.

Whereas the above commentators elaborated on how a goal might demonstrate selfishness, **Huebner & Rupert** suggest avoiding such terms altogether. To those authors, constructs such as goals, selfishness, and conscious versus unconscious processes are perhaps too deeply rooted in "traditional folk-taxonomies" because various forms of representation have motivational force. We see conceptual similarities between this view and other commentaries that address expanding the goal construct. For example, **Nanay** makes the distinction between specific and nonspecific behavioral tendencies; whereas **Eitam & Higgins**'s ROAR framework charts the ways in which a primed or activated latent tendency emerges into action only when a specific second condition is satisfied.

We disagree with **Huebner & Rupert** that building our model around the goal construct is "likely to inhibit progress in cognitive science." The goal construct may have roots in folk taxonomy, but it also has a deep and meaningful tradition in psychology (for reviews, see Austin & Vancouver 1996; Gollwitzer & Moskowitz 1996); the existence of this literature helps clarify specific mechanisms of goal selfishness and points to future directions of research. For example, the goal literature is intimately connected to concepts of reward, reinforcement, incentive, and affect not typically associated with cognitive representations (Kruglanski et al. 2002). (Indeed, as previously mentioned, many of the commentators rightly noted our relative neglect of affect as a variable that helps explain "why" people pursue certain end-states over others.)

Moreover, the goal construct also implicates notions of reference values, discrepancy reduction, and hierarchical feedback loops, which are central concepts within the control theory literature (Carver & Scheier 1982) and are likely to guide future ideas about pursuit. Consequently, we share **Pezzulo**'s belief that hierarchical control systems provide one of the most promising directions for future research in goal pursuit. For example, a model in which action becomes goal-directed because behaviors are steered in ways that minimize prediction errors between the goal (desired end-state) and the present (current state) is very much in harmony with Carver and

Scheier's (1982) seminal homeostatic model of goal pursuit via continual discrepancy reduction, and we elaborate upon this below.

R3. Goals in vivo

In the target article we argued that a single active goal can be considered "selfish" in reference to the person expressing it. Because that relation was our main focus, we included only a limited discussion of how goal conflict is resolved between goals, suggesting that a "winner take all" mechanism leads a single goal to dominate (sect. 3.2.3; similar to the "full-system" orientation mentioned in commentaries by **Becker & Kenrick**, **Müller & Amato**, and **Neuberg & Schaller**). Admittedly, the proposed mechanism may be most relevant for some goals (particularly powerful ones, such as evolutionarily adaptive motives, or addictions) as a result, we extend discussion of this topic here.

As the commentators recognized, there are many issues to ponder when considering mechanisms of goal conflict resolution, even in the initial stages of the task. To begin with, one should consider the broader body of research on motivational hierarchies and the specific instances where goals can be said to cooperate. Both topics were of particular emphasis in the commentaries, and here we suggest that cybernetic models of control (e.g., Perceptual Control Theory; Powers 1973) may help address these issues, as well as the open questions regarding goal dynamics.

Although the literature on control systems is vast and encompasses a variety of paradigms, it may help explain how relatively autonomous units (goals) give rise to complex systems capable of accommodating environmental variability (individual behavior). (For more comprehensive treatments of control systems within psychology, see Carver & Scheier [1998] and Powers [1973].)

R3.1. Control systems

Broadly stated, control-systems models explain how a comparatively simple system can produce reliable patterns of behavior (i.e., attain goals) in the seemingly infinitely variable environment of the real world. The basic control-system unit of behavioral organization has four main components: (1) an input function (also called a "perception"); (2) a standard (or "reference signal") against which the current state is compared; (3) comparison of the input against the standard that leads to action given sensed discrepancies; and (4) a feedback effect whereby the action changes the environment, thus updating the original perceptual signal (because actions make an aspect of the world come to a new state; Powers 1973). These components stand in relation to one another, and together they form a "loop" whereby organisms act to control how the environment is affecting them (in both positive and negative ways; also called "negative" or "discrepancy-reducing" feedback loops; Carver & Scheier 2002). Over time, it has the effect of reinforcing (if the environmental feedback is positive) or diminishing (if the environmental feedback is negative) certain behaviors.

Along with others (Carver & Scheier 2002, p. 305), we find this framework congenial for developing more systems-based understandings of individual behavior because people can be viewed as organizations of self-regulating feedback systems. Input functions represent the current state of the individual (within the environment, and vis-à-vis other goals and inputs from other systems levels). Desired end-states are the standard against which a person's current experienced state is compared. If the input function indicates discrepancies between actual and desired behavior (i.e., the goal has not yet been fulfilled), the individual executes behaviors to try to minimize the detected discrepancy until it is eliminated. A person's streamlined actions can be interpreted as the result of control systems making continual adjustments based upon the feedback (perceptions) these actions produce from the environment.

R3.2. (Self-) organization in goal systems

Within a control system, feedback loops can be linked hierarchically, with higher levels of control broadly correlating with the abstraction of the input (the input itself can span concrete events such as whether specific sensory nerve endings are stimulated or whether sequences of action can be understood according to an abstract standard such as honesty). Indeed, many commentators (e.g., **Mazzone**; **Pezzulo**; and **Sripada, Swain, Ho, & Swain** [**Sripada et al.**]) steered our attention to how factors at higher levels of processing influence pursuit of a focal goal.

A hierarchical organization of control has the potential to accommodate many of the abstract variables mentioned in the commentaries, including a motivational self **(Baumeister & Winegard**, **Fishbach**, and **Hirsh)** and the motivational relevance of representations (**Eitam & Higgins**). That also recalls **Nanay**'s distinction between *specific* versus *unspecific* goals, insofar as goals at multiple levels of abstraction also have significantly different means and contexts in which they can be successfully pursued (an intuition echoed by **Mazzone**). We add that higher- and lower-level goals may have different relationships to their goal-means. For example, one can avoid physical harm from a predator through a limited number of ways, including fighting, fleeing, or freezing. A higher-level goal, however, has the potential to be satisfied in a variety of more abstract ways and may implicate a wider breadth of processes to reconfigure. One can cooperate both by volunteering one's time on a survey or by picking up pens dropped by an experimenter. At an even higher level of abstraction, one can make one's parent proud by achieving in school, becoming rich, doing good deeds, or winning athletic competitions.

If they agreed on the importance of goal integration, the commentators disagreed about which goals are most central and how they assemble together. Many commentaries focused on the importance of conscious processes for the integration of constituent goals. For example, **Baumeister & Winegard** (see also **Fishbach**) proposed that traditional notions of a "conscious self" can be seen as consisting partly of processes that favor some goals over others. **Eitam & Higgins**'s commentary reviews how broader contextual factors related to control, relevance, and value similarly sway the relative goal priority. **Hirsh**'s commentary describes how effort and self-reflection are critical processes to the creation of a more coherent "self-system," whereas **Sripada et al.** provided thematically similar evidence from neuroscience, pointing out that the neuroregions traditionally involved with the resolution of goal competition are also implicated in higher-order processing and conscious self-regulation.

Conversely, **Becker & Kenrick** (as well as **Conroy-Beam & Buss** and **Neuberg & Schaller**) endorsed bottom-up processes that eventually create a hierarchical goal structure. Earlier-developing pursuits that are directly linked to adaptive outcomes (e.g., self-protection) serve as the foundation of a pyramid of universal human needs; they assume greatest priority for goal selection in proximal contexts. Goals that are higher up in the hierarchy (e.g., mate acquisition) develop later in life and are pursued only after successful pursuit of other, more foundational goals.

Carver and Scheier (2002) note that within control systems, integration can occur through relatively autonomous, bottom-up processes (endemic to the self-organization capabilities of dynamic computer systems), as well as through top-down influences (traditionally associated with the term *self-regulation*). That leads to the interesting prediction that the ways in which goal hierarchies are formed may result in different dynamics between the goals organized within that system. Below, we review evidence that various forms of relationships exist between goals in a system, and predict how these patterns might vary according to their place in the hierarchical structure.

R3.3. Selfish goals cooperate

The idea that goals exert reliably selfish influences on *the individual* should not imply that they are always in competition with *each other* and never surrender the steering wheel (with pathological and physiologically powerful cases such as addictions being a potential exception; see **Müller & Amato**). That point was not sufficiently highlighted in our target article, and because leaving this point unclarified would underestimate the scope of behavior that could be incorporated into our model, we spend some time discussing it here.

Thankfully, the commentators had similar intuitions about the relevance of the body of research on goal cooperation; their responses helped correct for its absence in the target article. We outline these below, but we also direct readers to **Kopetz, Hofmann, & Wiers** (**Kopetz et al.**), **Fishbach**, and **Carruthers** for better descriptions of the empirical evidence than have space for in this reply.

The Selfish Goal model holds that "goals often can and do encourage behaviors that are consistent with (or at least not opposed to) other goals' end-states" (sect. 4, para.1). Indeed, recent social cognitive findings support the notion that goal activation can result not only in selfish effects on the person, but also in cooperative effects that aid in the pursuit of other goals. As mentioned by **Kopetz et al.**, when a currently active goal and an alternate goal are perceived as unrelated to each other, priming the alternate goal undermines persistence and performance on the active goal (which can be seen as evidence for goals competing for the individual's limited cognitive resources). However, when goals are perceived to facilitate each other, priming the alternate goal actually increases persistence and performance on the focal goal (Shah et al. 2002) – a phenomenon that is simultaneously cooperative and self-interested.

Goals can also collaborate through more explicit forms of collusion, as when a person's expressed behaviors represent two different end-states (here *parallel constraint satisfaction* models are particularly relevant; see Kunda & Thagard 1996). **Carruthers** describes a notable example of goal collusion in a series of studies. When people are induced to write opinion essays (as in classic cognitive dissonance paradigms), the expression of the goal to say what one believes is constrained by the requirements of the goal of positive self-presentation. The outcome of the goal conflict is an expressed attitude that falls midway between the attitudes associated with pursuing each goal independently.

Another instance of cooperative goal behavior (or at least behavior that is interpretable in that light) occurs when a goal representation becomes inhibited following completion of the goal pursuit attempt (Förster et al. 2005). As Atkinson and Birch (1970) first argued, the function of the goal turn-off effect is to give other important goals their chances at attainment. **Fishbach** and **Kopetz et al.** significantly expand upon this idea, identifying cooperative behaviors in cases where goals are pursued in serial fashion and hence can be seen as cooperative. For example, the fulfillment of one goal may lead to the activation another, as when a person indulges in a chocolate dessert only if he or she has exercised earlier in the day, thereby "justifying" the current pursuit (i.e., changing its incentive value from negative to positive) through prior attainment of another goal.

If selfish goals can compete, operate simultaneously with another goal, or drive pursuit contingent on the completion of another goal, it remains to be seen what factors predict when these relational dynamics will occur. Future research may uncover different forms of goal cooperation and balancing depending on how (or when, in evolution or in development; see **Fedyk & Kushnir**) they were incorporated into higher-level structures. For example, the goal balancing phenomenon attributed to the operation of **Fishbach**'s motivational self may not characterize the behavior of lateral goals that are linked, bottom-up fashion, in **Becker & Kenrick's** oligarchy. Exercising may release a person to eat junk food, but successfully avoiding disease may not induce a person to become more amenable to other forms of physical danger.

Developmental stage may also offer insight into when these relational dynamics will occur (**Fedyk & Kushnir** and **Hirsh**). Goals that were integrated early in the pursuer's developmental stage may be systematically different than those incorporated in relatively later stages of adulthood, when conscious capabilities can assist with the task. One might expect, for example, that the goal facilitation as described by Shah et al. (2002) (also reviewed by **Fishbach** and **Kopetz et al.**) may be particularly likely to characterize goals that are integrated relatively late in development. One might additionally expect that the ability to use one means to fulfill two goals (as described in **Carruthers**'s commentary) becomes more effective with age and experience, as pursuers develop sophisticated ways to perceive how a single behavioral sequence can represent or satisfy two different end-states.

R3.4. The goal standards of society

Some commentators highlighted the societal implications that ought to be considered in light of the target article.

People's actions are significantly constrained by social norms and sanctions that in large measure determine the incentive and disincentive value of the various goal pursuits. Behaviors are observed and reconciled within the context of a greater society of individuals, and as described in **Washington & Kelly**'s commentary, the issues raised have implications for the moral institutions that uphold society. The commentary argues that if individuals engage in post hoc rationalizations (and the fact that the information can be used to inform future behavior) or use feedback from behaviors to guide future behavior, these acts suggest that at least some guidance-related subsystem of processes exists within a person that can be held accountable. Thus, even if people lack traditional, centralized psychological cores that can be held "blameable" for their actions, actors can nevertheless be held morally responsible.

Indeed, people have important social-identity relationships with the larger institutions to which they belong, and how they measure and maintain these self-defining relationships is affected by their temporary goals (e.g., threat avoidance, system justification; impression-management). Consequently, understanding the mechanisms of human goal pursuit offers additional insight regarding how people relate to other social institutions. Particularly strong cases of selfish goal influence, such as with drug addictions, may challenge individuals' standing as normative members of society (e.g., **Cuzen et al.** and **Müller & Amato**).

Research suggests that goals can affect how a person perceives his or her own place within those institutions, even in cases where the perceptions are self-defeating at the level of the individual person.

One particularly intriguing example of this counterintuitive goal influence is driven by people's need to believe that the institutional system they operate under is fair to them (**Kay & Jost**). Specifically, people have important goals to imbue their social, economic, and political systems with legitimacy, but often the unconscious pursuit of this goal ironically upholds the very institutional systems that may operate against their individual self-interests and thus oppress them. This dynamic results in a somewhat paradoxical effect in which the most socially disadvantaged members of society are the most likely to support the existing social system (e.g., people of low-income status are more likely to endorse the statement that economic inequality is legitimate and necessary, as compared to those of high-income status; Jost et al. 2003).

As **Kay & Jost** note, examining the general structure of goals (for example, their contextual dependence or their reconfiguring abilities over a person's perceptions and behaviors) offers advantages over a focus on goal content (i.e., comparing which of two goals is more primary) for practical reasons as well as theoretical reasons. Knowledge of goal structure can be applied to interventions that alter features of the environment to effect change (which can be a particularly effective intervention; e.g., Johnson & Goldstein 2003). We would only add that it is possible that environmental changes spur behavioral change in part because the intervention leverages the environmentally dependent and selfish quality of goals.

Goal operation can reconfigure one's very experience of the world – as, for example, altering how irresistible desserts appear to the would-be dieters who wish to avoid them, or emphasizing the proximity and instrumentality of some means for one's ends (even if those means are generally socially undesirable and the goal is temporary, as when violence is seen as a means for the status-maintenance goal; Griskevicius et al. 2009). Consequently, eliminating the presence of opportunity conditions for unwanted goals decreases the likelihood that those pursuits will be activated or successfully pursued because it impedes the chain of processes that might otherwise perpetuate the unwanted pursuit in a person's corpus of behavior (see above in sect. R3.1).

More important, an environmental change allows for the possibility that another goal may be activated or pursued in that situation; and if we have argued for one thing, it is that goals, when activated, will influence a person in a manner which can be meaningfully understood as selfish. Consequently, once an alternate, equally selfish yet socially valued goal is activated, given successful attainment of the end-state and appropriate circumstances, that goal will become stronger (i.e., its influence over how the individual may become more automatic and reliable).

R4. Conclusions

In closing, we return to some particular themes of the target article. In it we argued for the primacy of the currently active goal as an important influence over human social behavior. Importantly, we suggested that this was the case whether that goal was selected and put into motion by conscious, intentional means or unconsciously by features of the environment, thereby reserving any stance on the specific function of consciousness within goal pursuit.

Extensive recent research and theory on evolutionary social cognition, as well as on infantile motivations and innate social expectations and preferences (e.g., for fairness and equity, as well as in-group favoritism; see Baillargeon et al. 2013), has expanded the domain of unconscious processes. Importantly, this emerging research also points to new insights regarding how goals come to be capable of unconscious operation in the first place. The evolutionary social psychologists certainly appreciated this point, but other commentaries gave us the impression that an outdated impression of unconscious processes may linger, carrying with it the notion that all goal pursuits start out as conscious and intentional and only become unconscious after considerable experience (i.e., the skill acquisition model of nonconscious processes). A quarter century ago the process of skill acquisition, or sublimation with frequent and consistent experience, was the only way one of us saw clear to the possibility of unconscious goal pursuit (Bargh 1990). Since then, however, the considerable and significant advances regarding evolutionary and early childhood social cognition and motivation call for an updating of any remaining anachronistic, exclusively "conscious first" notions of automaticity and unconscious processes, across psychology.

Finally, in addition to the autonomy of active goal pursuit, we would like to highlight the remarkable *plasticity* of the cognitive and evaluative apparatus shown by the power of the active goal to reconfigure the mental system to facilitate goal attainment. Other social psychologists have recently remarked on this plasticity (e.g., Cunningham et al. 2008; Fiske 2013). It is shown, for example, in the research on alliance formation and how it changes stereotypic evaluations of an out-group member (which was long believed to be a chronic and intransigent tendency) from negative to positive if that out-group member suddenly becomes a teammate (Sherif et al. 1961). Alliances shift and friends become foes and vice versa. For the sake of adaptability to the shifting and changing social landscape, our goal pursuits can be most successful if the subprocesses they invoke are not slaves to our pasts. We believe such plasticity is an important component of autonomous goal operation, one that is also consistent with emerging research findings on epigenetics, the "affective forecasting" research showing how quickly we adapt to dramatic changes in our life situation, as well as continuing advances in knowledge on the adaptability of the human infant to the particulars of the social and physical environment in which it happened to be born.

In closing, we again express our appreciation for the opportunity to argue for the selfishness of goal pursuit, both in the target article and here in this response. The commentators' contributions challenged and helped us to engage with our model in new ways. Regardless of whether readers are ultimately convinced of our central argument, as long as they have considered existing research in social psychology from this new perspective, we will have attained our own (selfish) goals.

References

[The letters "a" and "r" before author's initials stand for target article and response references, respectively]

Aarts, H. (2007) On the emergence of human goal pursuit: The nonconscious regulation and motivation of goals. *Social and Personality Psychology Compass* 1(1):183–201. [EB-M]

Aarts, H., Custwers, R. & Holland, R.W. (2007) The nonconscious cessation of goal pursuit: When goals and negative affect are coactivated. *Journal of Personality and Social Psychology* 92:165–78. Available at: http://goallab.nl/publications/documents/Aarts,%20Custers,%20Holland%20%282007%29%20-%20goal%20priming%20and%20negative%20affect.pdf. [aJYH, GBM]

Aarts, H. & Dijksterhuis, A. (2000) Habits as knowledge structures: Automaticity in goal-directed behavior. *Journal of Personality and Social Psychology* 78:53–63. Available at: http://goallab.nl/publications/documents/Aarts,%20Dijksterhuis%20%282000%29%20-%20habits%20as%20knowlegde%20structures.pdf. [aJYH]

Ackerman, J. M., Becker, D. V., Mortensen, C. R., Sasaki, T., Neuberg, S. L. & Kenrick, D. T. (2009) A pox on the mind: Disjunction of attention and memory in the processing of physical disfigurement. *Journal of Experimental Social Psychology* 45:478–85. [DVB, SLN]

Agassi, J. (1968) Antropomorphism in science. In: *Dictionary of the history of ideas: Studies of selected pivotal ideas*, ed. P. P. Wiener, pp. 87–91. Scribner. Available at: http://www.tau.ac.il/~agass/joseph-papers/anthro.pdf. [TAM]

Ainslie, G. (1992) *Picoeconomics: The strategic interaction of successive motivational states within the person*. Cambridge University Press. [GA, DS]

Ainslie, G. (2007) Can thought experiments prove anything about the will? In: *Distributed cognition and the will: Individual volition and social context*, ed. D. Spurrett, D. Ross, H. Kincaid & L. Stephens, pp. 169–96. MIT. [GA]

Ainslie, G. (2012) Pure hyperbolic discount curves predict "eyes open" self-control. *Theory and Decision* 73:3–34. doi: 10.1007/s11238-011-9272-5. [GA]

Ainslie, G. (2013) Grasping the impalpable: The role of endogenous reward in process addictions. *Inquiry* 56:446–69. DOI: 10.1080/0020174X.2013.806129. Avialable at: http://www.tandfonline.com/eprint/8fGTuFsnfFunYJKJ7aA7/full.. [GA]

Ajzen, I. & Fishbein, M. (1980) *Understanding attitudes and predicting social behavior*. Prentice-Hall. [aJYH]

Akins, K. (1996) Of sensory systems and the "aboutness" of mental states. *Journal of Philosophy* 93(7):337–72. [BH]

Alcock, J. (2001) *The triumph of sociobiology*. Oxford University Press. [MF]

Allman, J. M. (2000) *Evolving brains*. Scientific American Library. [aJYH]

Allport, G. W. (1937) *Personality: A psychological interpretation*. Holt. [JBH]

Amburgey, T. L. & Rao, H. (1996) Organizational ecology: Past, present, and future directions. *Academy of Management Journal* 39:1265–86. Available at: http://www.jstor.org/stable/256999. [rJYH]

American Psychiatric Association (1994) *Diagnostic statistical manual of mental disorders, 4th ed. (DSM-IV)*. American Psychiatric Association. [CPM]

Anderson, E., Siegel, E. H. & Barrett, L. F. (2011a) What you feel influences what you see: The role of affective feelings in resolving binocular rivalry. *Journal of Experimental Social Psychology* 47(4):856–60. [EB-M]

Anderson, E., Siegel, E. H., Bliss-Moreau, E. & Barrett, L. F. (2011b) The visual impact of gossip. *Science* 332(6036):1446–48. [EB-M]

Anderson, M. L. (2010) Neural reuse: A fundamental organizational principle of the brain. *Behavioral and Brain Sciences* 33:245–66. Available at: http://www.agcognition.org/papers/anderson_bbs_2010.pdf. [aJYH]

Anderson, N. H. (1968) Likableness ratings of 555 personality-trait words. *Journal of Personality and Social Psychology* 9:272–79. [aJYH]

Anderson, R. C. & Pichert, J. W. (1978) Recall of previously unrecallable information following a shift in perspective. *Journal of Verbal Learning and Verbal Behavior* 17:1–12. Available at: http://homepage.psy.utexas.edu/homepage/Faculty/Markman/PSY394/Anderson_Pichert.pdf. [aJYH]

Aronson, E. (1995) *The social animal*. Worth. [aJYH]

Asch, S. E. (1952) *Social psychology*. Prentice-Hall. [aJYH]

Atkinson, J. W. (1964) *An introduction to motivation*. Van Nostrand. [CK]

Atkinson, J. W. & Birch, D. (1970) *The dynamics of action*. Wiley. [arJYH]

Aunger, R. & Curtis, V. (2013) The anatomy of motivation: An evolutionary-ecological approach. *Biological Theory* 8:49–63. [SLN]

Austin, J. T. & Vancouver, J. B. (1996) Goal constructs in psychology: Structure, process, and content. *Psychological Bulletin* 120:338–75. [arJYH, JBH]

Baars, B. J. (1986) *The cognitive revolution in psychology*. Guilford. [GA]

Baddeley, A. D. (2003) Working memory: Looking back and looking forward. *Nature Reviews Neuroscience* 4:829–39. Available at: http://www.nature.com/nrn/journal/v4/n10/abs/nrn1201.html. [aJYH]

Baddeley, A. D. & Hitch, G. (1974) Working memory. In: *The psychology of learning and motivation, Vol. 8*, ed. G. H. Bower, pp. 47–89. Academic Press. [aJYH]

Bagozzi, R. P. & Pieters, R. (1998) Goal-directed emotions. *Cognition & Emotion* 12 (1):1–26. [EB-M]

Baillargeon, R., He, Z., Setoh, P., Scott, R.M., Sloan, S. & Yan D.Y. (2013) False-belief understanding and why it matters: The social-acting hypothesis. In: *Navigating the social world: What infants, children, and other species can teach us*, ed. M. Banaji & S. Gelmanm, pp. 88–95. Oxford University. [rJYH]

Baker, T. B., Piper, M. E., McCarthy, D. E., Majeskie, M. R. & Fiore, M. C. (2004) Addiction motivation reformulated: An affective processing model of negative reinforcement. *Psychological Review* 111:33–51. Available at: http://dionysus.psych.wisc.edu/lit/Articles/BakerT2004a.pdf. [aJYH]

Balcetis, E. & Dunning, D. (2006) See what you want to see: Motivational influences on perception. *Journal of Personality and Social Psychology* 91:612–25. [aJYH, GBM]

Balcetis, E. & Dunning, D. (2010) Wishful seeing: Desired objects are seen as closer. *Psychological Science* 21:147–52. [aJYH, GBM]

Balcetis, E., Dunning, D. & Granot, Y. (2012) Subjective value determines initial dominance in binocular rivalry. *Journal of Experimental Social Psychology* 48:122–29. [GBM]

Balleine, B. W. & O'Doherty, J. P. (2010) Human and rodent homologies in action control: Corticostriatal determinants of goal-directed and habitual action. *Neuropsychopharmacology Review* 35:48–69. [NLC]

Band, G. P., van Steenbergen, H., Ridderinkhof, K. R., Falkenstein, M. & Hommel, B. (2009) Action-effect negativity: Irrelevant action effects are monitored like relevant feedback. *Biological Psychology* 82(3):211–18. [BE]

Bandura, A. (1977) Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review* 84:191–215. Available at: http://des.emory.edu/mfp/Bandura1977PR.pdf. [arJYH]

Bandura, A. (1986) Social foundations of thought and action: A social cognitive theory. Prentice-Hall. [aJYH]

Bandura, A. (1991) Self-regulation of motivation through anticipatory and self-reactive mechanisms. In: *Nebraska symposium on motivation, 1990*, ed. R. A. Dienstbier, pp. 69–164. University of Nebraska Press. [AF]

Bar, M. (2007) The proactive brain: Using analogies and associations to generate predictions. *Trends in Cognitive Sciences* 11:280–89. [GP]

Bar-Anan, Y., Wilson, T. D. & Hassin, R. R. (2010) Inaccurate self-knowledge formation as a result of automatic behavior. *Journal of Experimental Social Psychology* 46:884–94. [arJYH]

Barch, D. M., Braver, T. S., Akbudak, E., Conturo, T., Ollinger, J. & Snyder, A. (2001) Anterior cingulate cortex and response conflict: Effects of response modality and processing domain. *Cerebral Cortex* 11(9):837–48. Available at: http://www.ncbi.nlm.nih.gov/pubmed/11532889. [CS]

Bargh, J. A. (1984) Automatic and controlled processing of social information. In: *Handbook of social cognition, Vol. 1*, ed. R. S. Wyer Jr. & T. K. Srull, pp. 1–41. Erlbaum. [aJYH]

Bargh, J. A. (1989) Conditional automaticity: Varieties of automatic influence on social perception and cognition. In: *Unintended thought*, ed. J. Uleman & J. Bargh, pp. 3–51. Guilford. Available at: http://www.yale.edu/acmelab/articles/Bargh1989.pdf. [aJYH]

Bargh, J. A. (1990) Auto-motives: Preconscious determinants of social interaction. In: *Handbook of motivation and cognition, vol. 2*, ed. E. T. Higgins & R. M. Sorrentino, pp. 93–130. Guilford. Available at: http://www.yale.edu/acmelab/articles/Bargh_1990.pdf. [arJYH, GBM]

Bargh, J. A. (1992) Does subliminality matter to social psychology: Awareness of the stimulus versus awareness of its influence. In: *Perception without awareness*, ed. R. F. Bornstein & T. S. Pittman, pp. 236–55. Guilford. [aJYH]

Bargh, J. A. (1994) The four horsemen of automaticity: Awareness, intention, efficiency, and control in social cognition. In: *Handbook of social cognition, 2nd ed.*, ed. R. S. Wyer, Jr. & T. K. Srull, pp. 1–40. Erlbaum. [aJYH]

Bargh, J. A. (1997) The automaticity of everyday life. In: *Advances in social cognition, vol. 10,* ed. R. S. Wyer, Jr., pp. 1–61. Erlbaum. [GA]

Bargh, J. A. (1999) The cognitive monster: The case against controllability of automatic stereotype effects. In: *Dual process theories in social psychology*, ed. S. Chaiken & Y. Trope, pp. 361–82. Guilford. Available at: http://www.yale.edu/acmelab/articles/Bargh_1999_Cog_Monster.pdf. [aJYH]

Bargh, J.A. (2006) What have we been priming all these years? On the development, mechanisms, and ecology of nonconscious social behavior. *European Journal of Social Psychology* 36:147–68. Available at: http://www.yale.edu/acmelab/articles/Bargh_EJSP_2006.pdf. [aJYH, GA]

Bargh, J. A., ed. (2007) *Social psychology and the unconscious: The automaticity of higher mental processes*. Psychology Press. [aJYH]

Bargh, J. A. & Barndollar, K. (1996) Automaticity in action: The unconscious as repository of chronic goals and motives. In: *The psychology of action: Linking cognition and motivation to behavior*, ed. P. Gollwitzer & J. A. Bargh, pp. 457–81. Guilford. [JBH]

Bargh, J. A. & Chartrand, T. L. (2000) The mind in the middle: A practical guide to priming and automaticity research. In: *Handbook of research methods in social and personality psychology*, ed. H. T. Reis & C. M. Judd, pp. 253–316. Cambridge University Press. [aJYH]

Bargh, J. A., Chen, M. & Burrows, L. (1996) Automaticity of social behavior: Direct effects of trait construct and stereotype priming on action. *Journal of Personality and Social Psychology* 71:230–44. Available at: http://www.yale.edu/acmelab/articles/Chen_Bargh_PSPB1999.pdf. [aJYH]

Bargh, J. A. & Ferguson, M. J. (2000) Beyond behaviorism: The automaticity of higher mental processes. *Psychological Bulletin* 126:925–45. [aJYH]

Bargh, J. A., Gollwitzer, P. M., Lee-Chai, A., Barndollar, K. & Trotschel, R. (2001) The automated will: Unconscious activation and pursuit of behavioral goals. *Journal of Personality and Social Psychology* 81:1004–27. Available at: http://www.yale.edu/acmelab/articles/AutomatedWill2001.pdf. [aJYH, GBM]

Bargh, J. A., Gollwitzer, P. M. & Oettingen, G. (2010) Motivation. In: *Handbook of social psychology* (5th ed.), ed. S. Fiske, D. Gilbert & G. Lindzey, pp. 268–316. Wiley. [aJYH]

Bargh, J. A., Green, M. & Fitzsimons, G. (2008) The selfish goal: Unintended consequences of intended goal pursuits. *Social Cognition* 26:520–40. Available at: http://www.yale.edu/acmelab/articles/SelfishGoal_SocialCognition.pdf. [arJYH]

Bargh, J. A. & Morsella, E. (2008) The unconscious mind. *Perspectives on Psychological Science* 3:73–9. Available at: http://www.yale.edu/acmelab/articles/Bargh_Morsella_Unconscious_Mind.pdf. [aJYH]

Bargh, J. A. & Morsella, E. (2010) Unconscious behavioral guidance systems. In: *Then a miracle occurs: Focusing on behavior in social psychological theory and research,* ed. C. Agnew, D. Carlston, W. Graziano & J. Kelly, pp. 89–118. Oxford University Press. [aJYH]

Bargh, J. A., Schwader, K. L., Hailey, S. E., Dyer, R. L. & Boothby, E. J. (2012) Automaticity in social-cognitive processes. *Trends in Cognitive Science* 16:593–605. [aJYH]

Barnes, J., Dong, C. Y., McRobbie, H., Walker, N., Mehta, M. & Stead, L. F. (2010) Hypnotherapy for smoking cessation. *Cochrane Database System Review* 2010; (10): CD001008. [GA]

Barrett, L. F. (2006) Solving the emotion paradox: Categorization and the experience of emotion, *Personality and Social Psychology Review* 10(1):20–46. [EB-M]

Barrett, L. F. & Bliss-Moreau, E. (2009) Affect as a psychological primitive. *Advances in Experimental Social Psychology* 41:167–218. [EB-M]

Barrett, L. F. & Russell, J. A. (1999) The structure of current affect: Controversies and emerging consensus. *Current Directions in Psychological Science* 8(1):10–14. [EB-M]

Baumeister, R. F., Bratslavsky, E., Muraven, M. & Tice, D. M. (1998) Ego depletion: Is the active self a limited resource? *Journal of Personality and Social Psychology* 74:1252–65. Available at: http://faculty.washington.edu/jdb/345/345%20Articles/Baumeister%20et%20al.%20(1998).pdf. [AF, arJYH]

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* 117:945–71. Available at: http://www.psy.fsu.edu/~masicampo/BaumeisterMasicampo2010.pdf. [aJYH]

Baumeister, R. F., Schmeichel, B. J. & Vohs, K. D. (2007) Self-regulation and the executive function: The self as controlling agent. In: *Social psychology: Handbook of basic principles, 2nd ed.*, ed. A. W. Kruglanski & E. T. Higgins, pp. 516–39. Guilford. Available at: https://www.csom.umn.edu/Assets/71708.pdf. [rJYH]

Becker, D. V., Anderson, U. S., Neuberg, S. L., Maner, J. K., Shapiro, J. R., Ackerman, J. M., Schaller, M. & Kenrick, D. T. (2010) More memory bang for the attentional buck: Self-protection goals enhance encoding efficiency for potentially threatening males. *Social Psychological and Personality Science* 1:182–89. [DVB]

Becker, D. V., Mortensen, C. R., Ackerman, J. M., Shapiro, J. R., Anderson, U. S., Sasaki, T., Maner, J. K., Neuberg, S. L. & Kenrick, D. T. (2011) Signal detection on the battlefield: Priming self-protection vs. revenge-mindedness differentially modulates the detection of enemies and allies. *PLoSOne*, 6, e23929. doi: 10.1371/journal.pone.0023929. [DVB]

Bélanger, J. J., Lafrenière, M. K., Vallerand, R. J. & Kruglanski, A. W. (2013) When passion makes the heart grow colder: The role of passion in alternative goal suppression. *Journal of Personality and Social Psychology* 104:126–47. [CK]

Belin, D. & Everitt, B. J. (2008) Cocaine seeking habits depend upon dopamine-dependent serial connectivity linking the ventral with the dorsal striatum. *Neuron* 57(3):432–41. [CPM]

Belin, D., Jonkman, S., Dickinson, A., Robbins, T. W. & Everitt, B. J. (2009) Parallel and interactive learning processes within the basal ganglia: Relevance for the understanding of addiction. *Behavioural Brain Research* 199:89–102. [CPM]

Bem, D. (1967) Self-perception: An alternative interpretation of cognitive dissonance phenomena. *Psychological Review* 74:183–200. [PC]

Bem, D. J. (1972) Self-perception theory. In: *Advances in experimental social psychology, vol. 6,* ed. L. Berkowitz, pp. 1–62. Academic Press. [AF]

Bernard, L. C., Mills, M., Swenson, L. & Walsh, R. P. (2005) An evolutionary theory of human motivation. *Genetic, Social, and General Psychology Monographs* 131:129–84. [SLN]

Berns, G. S., Laibson, D. & Loewenstein, G. (2007) Intertemporal choice –Toward an integrative framework. *Neuroeconomics* 11:482–88. [GA]

Bliss-Moreau, E. & Barrett, L. F. (2009) What's reason got to do with it? Affect as the foundation of learning. *Behavioral and Brain Sciences* 32(2):201–202. [EB-M]

Bogen, J. E. (1995) On the neurophysiology of consciousness: II. Constraining the semantic problem. *Consciousness and Cognition* 4:137–58. Available at: http://www.ncbi.nlm.nih.gov/pubmed/8521253. [aJYH]

Bongers, K. C. A., Dijksterhuis, A. & Spears, R. (2009) Self-esteem regulation after success or failure to attain unconsciously activated goals. *Journal of Experimental Social Psychology* 45:468–77. [aJYH]

Bor-Sen, C., Chih-Yuan, H. & Jing-Jia, L. (2011) Robust design of biological circuits: Evolutionary systems biology approach. *Journal of Biomedicine and Biotechnology* 2011:304236. [TAM]

Botvinick, M. M., Braver, T. S., Barch, D. M., Carter, C. S. & Cohen, J. D. (2001) Conflict monitoring and cognitive control. *Psychological Review* 108(3):624–52. Available at: http://www.ncbi.nlm.nih.gov/pubmed/11488380. [CS]

Botvinick, M. M., Cohen, J. D. & Carter, C. S. (2004) Conflict monitoring and anterior cingulate cortex: An update. *Trends in Cognitive Science* 8(12):539–46. Available at: http://www.ncbi.nlm.nih.gov/pubmed/15556023. [CS]

Brandt, M. J. (2013) Do the disadvantaged legitimize the social system? A large-scale test of the status-legitimacy hypothesis. *Journal of Personality and Social Psychology* 104:765–85. [ACK]

Brehm, J. W. (1966) *A theory of psychological reactance.* Academic Press. [ACK]

Brewer, M. B. (1988) A dual process model of impression formation. In: *Advances in social cognition Vol. 1*, ed. T. K. Srull & R. S. Wyer, Jr., pp. 1–36. Erlbaum. [aJYH]

Brill, A. A. (1938) Introduction. In: *The basic writings of Sigmund Freud,* ed. & trans. A. A. Brill, pp. 1–32. Modern Library. [aJYH]

Bruner, J. S. (1957) On perceptual readiness. *Psychological Review* 64:123–52. [BE, aJYH]

Buchsbaum, B. & D'Esposito, M. (2008) The search for the phonological store: From loop to convolution. *Journal of Cognitive Neuroscience* 20:762–78. [aJYH]

Bugental, D. B. (2000) Acquisition of the algorithms of social life: A domain-based approach. *Psychological Bulletin* 126:187–19. [aJYH]

Burt, A. & Trivers, R. (2006) *Genes in conflict: The biology of selfish genetic elements*. Harvard University Press. [WvH]

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

Buss, D. M. (2012) *Evolutionary psychology: The new science of the mind, 4th ed.* Allyn & Bacon. [DC-B]

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

53:532–48. Available at: http://homepage.psy.utexas.edu/homepage/Group/BussLAB/pdffiles/Adaptations_Exaptations_Spandrels_1998.pdf. [aJYH]

Buss, D. M., Larsen, R. J., Westen, D. & Semmelroth, J. (1992) Sex differences in jealousy: Evolution, physiology, and psychology. *Psychological Science* 3 (4):251–55. [DC-B]

Cabanac, M., Cabanac, A. J. & Parent, A. (2009) The emergence of consciousness in phylogeny. *Behavioural Brain Research* 198:267–72. doi: 10.1016/j.bbr.2008.11.028. [TAM]

Campbell, D. T. (1974) Evolutionary epistemology. In: *The philosophy of Karl Popper,* ed. P. A. Schilpp, pp. 413–63. Open Court Publishing. [aJYH]

Cantor, N. & Blanton, H. (1996) Effortful pursuit of personal goals in daily life. In: *The psychology of action: Linking cognition and motivation to behavior*, ed. J. A. Bargh & P. M. Gollwitzer, pp. 338–59. Guilford. [aJYH]

Carruthers, P. (2011) *The opacity of mind: An integrative theory of self-knowledge*. Oxford University Press. [PC]

Carter, C. S., Braver, T. S., Barch, D. M., Botvinick, M. M., Noll, D. & Cohen, J. D. (1998) Anterior cingulate cortex, error detection, and the online monitoring of performance. *Science* 280(5364): 747–49. Available at: http://www.ncbi.nlm.nih.gov/pubmed/9563953. [CS]

Carver, C. S. & Scheier, M. F. (1982) Control theory: A useful conceptual framework for personality-social, clinical, and health psychology. *Psychological Bulletin* 92:111–35. [rJYH]

Carver, C. S. & Scheier, M. F. (1990) Origins and functions of positive and negative affect: A control-process view. *Psychological Review* 97(1):19–35. [EB-M]

Carver, C. S. & Scheier, M. F. (1998) *On the self-regulation of behavior*. Cambridge University Press. [AF, JBH]

Carver, C. S. & Scheier, M. F. (1998) *On the self-regulation of behavior*. Cambridge University Press. [rJYH]

Carver, C. S. & Scheier, M. F. (2002) Control processes and self-organization as complementary principles underlying behavior. *Personality and Social Psychology Review* 6:304–15. [rJYH]

Cesario, J., Plaks, J. E. & Higgins, E. T. (2006) Automatic social behavior as motivated preparation to interact. *Journal of Personality and Social Psychology* 90:893–910. [GBM]

Chaiken, S. & Trope, Y. (1999) *Dual process theories in social psychology*. Guilford. [aJYH, CS]

Chartrand, T. L. & Bargh, J. A. (1996) Automatic activation of social information processing goals: Nonconscious priming reproduces effects of explicit conscious instructions. *Journal of Personality and Social Psychology* 71:464–78. Available at: http://www.yale.edu/acmelab/articles/chartrand_bargh_1996.pdf. [aJYH]

Chen, M. & Bargh, J. A. (1999) Consequences of automatic evaluation: Immediate behavioral predispositions to approach or avoid the stimulus. *Personality and Social Psychology Bulletin* 25:215–24. Available at: http://www.yale.edu/acmelab/articles/Chen_Bargh_PSPB1999.pdf. [aJYH]

Chen, S., Lee-Chai, A. Y. & Bargh, J. A. (2001) Relationship orientation as a moderator of the effects of social power. *Journal of Personality and Social Psychology* 80:173–87. Available at: http://www.yale.edu/acmelab/articles/chen_lee-chai_bargh_2001.pdf. [aJYH]

Chun, W. Y., Kruglanski, A. W., Friedman, R. & Sleeth-Keppler, D. (2011) Multifinality in unconscious choice. *Journal of Personality and Social Psychology* 101:1124–37. [GBM]

Cialdini, R. B., Schaller, M., Houlihan, D., Arps, K., Fultz, J. & Beaman, A. L. (1987) Empathy-based helping: Is it selflessly or selfishly motivated? *Journal of Personality and Social Psychology* 52:749–58. [GBM]

Cisek, P. & Kalaska, J. F. (2010) Neural mechanisms for interacting with a world full of action choices. *Annual Review of Neuroscience* 33:269–98. [GP]

Clark, A. (1997) *Being there: Putting brain, body, and world together again*. MIT Press. [BH]

Clark, A. (2008) Soft selves and ecological control, In: *Distributed cognition and the will,* ed. D. Spurrett, D. Ross, H. Kincaid & L. Stephens, pp. 101–22. MIT Press. [NW]

Clark, M. S. & Mills, J. (1993) The difference between communal and exchange relationships: What it is and is not. *Personality and Social Psychology Bulletin* 19:684–91. Available at: http://psp.sagepub.com/cgi/reprint/19/6/684. [aJYH]

Cohen, D., Nisbett, R. E., Bowdle, B. F. & Schwarz, N. (1996) Insult, aggression, and the Southern culture of honor: An "experimental ethnography." *Journal of Personality and Social Psychology* 70:943–60. [rJYH]

Cohen, J. D., Dunbar, K. & McClelland, J. L. (1990) On the control of automatic processes: A parallel distributed processing account of the Stroop effect. *Psychological Review* 97:332–61. Available at: http://www.csbmb.princeton.edu/ncc/publications/1990/CohenDunbarMcClelland1990.pdf. [aJYH]

Cole, S., Balcetis, E. & Dunning, D. (2013) Affective signals of threat produce perceived proximity. *Psychological Science* 24:34–40. [GBM]

Corballis, M. C. (2007) The evolution of consciousness. In: *The Cambridge handbook of consciousness*, ed. P. D. Zelazo, M. Moscovitch & E. Thompson, pp. 571–95. Cambridge University Press. [aJYH]

Corder, E. H., Saunders, A. M., Strittmatter, W. J., Schmechel, D. E., Gaskell, P. C., Small, G. W., Roses, A. D., Haines, J. L. & Pericak-Vance, M. A. (1993) Gene dose of apolipoprotein E type 4 allele and the risk of Alzheimer's disease in late onset families. *Science* 261:921–23. [WvH]

Cornelissen, J. P. (2004) Beyond compare: Metaphor in organization theory. *The Academy of Management Review* 30:751–64. [rJYH]

Cosmides, L. & Tooby, J. (2000) Evolutionary psychology and the emotions. In: *Handbook of emotions*, 2nd edition, ed. M. Lewis & J. M. Haviland-Jones. pp. 91–115. Guilford. [EB-M]

Cosmides, L. & Tooby, J. (2013) Evolutionary psychology: New perspectives on cognition and motivation. *Annual Review of Psychology* 64:201–09. [aJYH]

Crick, F. (1970) Central dogma of molecular biology. *Nature* 227:561–63. [TAM]

Crocker, J. & Park, L. E. (2004) The costly pursuit of self-esteem. *Psychological Bulletin* 130:392–414. Available at: http://wings.buffalo.edu/psychology/labs/SMRL/docs/Crocker%20and%20Park%20(2004a,%20Psych%20Bull).pdf. [aJYH]

Cunningham, W. A., Van Bavel, J. J. & Johnsen, I. (2008) Affective flexibility: Evaluative processing goals shape amygdala activity. *Psychological Science* 19:152–60. [rJYH]

Custers, R. & Aarts, H. (2005) Positive affect as implicit motivator: On the nonconscious operation of behavioral goals. *Journal of Personality and Social Psychology* 89:129–42. Available at: http://goallab.nl/publications/documents/Custers,%20Aarts%20%282005%29%20- %20positive%20affect%20as%20implicit%20motivator.pdf. [arJYH, BE]

Custers, R. & Aarts, H. (2010) The unconscious will: How the pursuit of goals operates outside of conscious awareness. *Science* 329:47–50. Available at: http://www.goallab.nl/publications/documents/Custers%20and%20Aarts%20-%20The%20unconscious%20will%20-%20Science%202010.pdf. [aJYH, GBM]

Custers, R., Maas, M., Wildenbeest, M. & Aarts, H. (2008) Nonconscious goal pursuit and the surmounting of physical and social obstacles. *European Journal of Social Psychology* 38:1013–22. Available at: http://www.goallab.nl/publications/documents/Custers,%20Maas,%20Wildenbeest,%20Aarts%20(2008)%20-%20nonconscious%20goal%20pursuit%20and%20obstacles.pdf. [aJYH]

da Fonseca, R. R., Kosiol, C., Vinar, T., Siepel, A. & Nielsen, R. (2010) Positive selection on apoptosis related genes. *FEBS Letters* 584:469–76. [TAM]

Daly, M., Wilson, M. I. & Weghorst, S. J. (1982) Male sexual jealousy. *Ethology & Sociobiology* 3:11–27. [DC-B]

Damasio, A. (1994) *Descartes' error*. Papermac. [PC]

Damon, W. & Hart, D. (1982) The development of self-understanding from infancy through adolescence. *Child Development* 53(4):841–64. [JBH]

Darley, J. & Latane, B. (1968) Bystander intervention in emergencies: Diffusion of responsibility. *Journal of Personality and Social Psychology* 8:377–83. Available at: http://faculty.uncfsu.edu/tvancantfort/Syllabi/Gresearch/Readings/A_Darley.pdf. [aJYH]

Darwin, C. (1859) *The origin of species*. P. F. Collier. [aJYH]

Darwin, C. R. (1871) *The descent of man, and selection in relation to sex*. John Murray. [TAM]

Daw, N. D. & Doya, K. (2006) The computational neurobiology of learning and reward. *Current Opinion in Neurobiology* 16:199–204. [GA]

Dawkins, R. (1976) *The selfish gene*. Oxford University Press. [arJYH, CPM, DS, WvH]

Dawkins, R. (1982) *The extended phenotype*. W. H. Freeman. [TAM]

Dayan, P. & Balleine, B. W. (2002) Reward, motivation, and reinforcement learning. *Neuron* 36:285–98. [GP]

de Wit, S. & Dickinson, A. (2009) Associative theories of goal-directed behaviour: A case for animal–human translational models. *Psychological Research* 73:463–76. [NLC]

de Wit, S., Watson, P., Harsay, H. A., Cohen, M. X., van de Vijver, I. & Ridderinkhof, K. R. (2012) Corticostriatal connectivity underlies individual differences in the balance between habitual and goal-directed action control. *Journal of Neuroscience* 32:12066–75. [NLC]

Deacon, T. W. (1997) *The symbolic species: The co-evolution of language and the brain*. W.W. Norton. [aJYH]

Dehaene, S. & Naccache, L. (2001) Towards a cognitive neuroscience of consciousness: Basic evidence and a workspace framework. *Cognition* 79:1–37. Available at: http://www.ncbi.nlm.nih.gov/pubmed/11164022. [aJYH]

Dehaene, S., Naccache, L., Le Clec`H, G., Koechlin, E., Mueller, M., Dehaene-Lambertz, G., van de Moortele, P. F. & Le Bihan, D. (1998) Imaging unconscious semantic priming. *Nature* 395:597–600. [BN]

Dennett, D. C. (1991) *Consciousness explained*. Little, Brown. [aJYH]

Dennett, D. C. (1995) *Darwin's dangerous idea: Evolution and the meanings of life*. Simon & Schuster. [aJYH]

Dennett, D. C. (2003) *Freedom evolves*. Penguin. [NW]

Desimone, R. & Duncan, J. (1995) Neural mechanisms of selective visual attention. *Annual Review of Neuroscience* 18:193–222. [GP]

Devine, P. G. (1989) Stereotypes and prejudice: Their automatic and controlled components. *Journal of Personality and Social Psychology* 56:5–18. Available at: https://mywebspace.wisc.edu/wtcox/web/trishpubs_files/Devine%20%281989%29.pdf?uniq=-c4okl9. [aJYH]

Dickinson, A. (1985) Actions and habits: The development of behavioural autonomy. *Philosophical Transactions of the Royal Society B: Biological Sciences* 308:67–78. [NLC]

Dijksterhuis, A. & Aarts, H. (2011) Goals, attention, and (un)consciousness. *Annual Review of Psychology* 61:467–90. [aJYH]

Dijksterhuis, A., Chartrand, T. L. & Aarts, H. (2007) Effects of priming and perception on social behavior and goal pursuit. In: *Social psychology and the unconscious: The automaticity of higher mental processes,* ed. J. A. Bargh, pp. 51–131. Psychology Press. [aJYH]

Dijksterhuis, A. & van Knippenberg, A. (1998) The relation between perception and behavior or how to win a game of Trivial Pursuit. *Journal of Personality and Social Psychology* 74:865–77. Available at: http://www.unconsciouslab.nl/publications/Dijksterhuis%20Van%20Knippenberg%20-%20The%20Relation%20Between%20Perception%20and%20Behavior.pdf. [aJYH]

Donald, M. (1991) *Origins of the modern mind*. Harvard University Press. [aJYH]

Donald, M. (1995) The neurobiology of human consciousness: An evolutionary approach. *Neuropsychologia* 33:1087–102. [TAM]

Doris, J. (in press) *Talking to ourselves: Reflection, skepticism, and agency*. Oxford University Press. [NW]

Dundzila, V. L. (1987) The ancient Latvian religion. *Lithuaninan Quarterly Journal of Arts and Sciences* 33:3. [TAM]

Dweck, C. S. & Leggett, E. L. (1988) A social-cognitive approach to motivation and personality. *Psychological Review* 95:256–73. [AF]

Echterhoff, G., Higgins, E. T. & Groll, S. (2005) Audience-tuning effects on memory: The role of shared reality. *Journal of Personality and Social Psychology* 89:257–76. [BE]

Effron, D. A., Cameron, J. S. & Monin, B. (2009) Endorsing Obama licenses favoring whites. *Journal of Experimental Social Psychology* 45:590–93. Available at: http://www-psych.stanford.edu/~monin/. [aJYH]

Eitam, B., Glass, R. S., Aviezer, H., Dienes, Z. & Higgins, E. T. (2013) Implicit learning of task irrelevant faces. Unpublished manuscript. [BE]

Eitam, B. & Higgins, E. T. (2010) Motivation in mental accessibility: Relevance of a representation (ROAR) as a new framework. *Social and Personality Psychology Compass* 4:951–67. [BE]

Eitam, B., Miele, D. B. & Higgins, E. T. (2013) Motivated remembering: Remembering as accessibility and accessibility as motivational relevance. In: *Handbook of social cognition*, ed. D. Carlston. Oxford University Press. [BE]

Elliot, A. & Devine, P. (1994) On the motivational nature of cognitive dissonance: Dissonance as psychological discomfort. *Journal of Personality and Social Psychology* 67:382–94. [PC]

Elliot, A. J. & Fryer, J. W. (2008) The goal construct in psychology. In: *Handbook of motivation science, vol. 18*, ed. J. Y. Shah & W. L. Gardner, pp. 235–50. Guilford. [aJYH]

Elster, J. (1990) Selfishness and altruism. In: *Beyond self-interest*, ed. J. J. Mansbridge, pp. 44–52. University of Chicago Press. [aJYH]

Emmons, R. A. (1986) Personal strivings: An approach to personality and subjective well-being. *Journal of Personality and Social Psychology* 51(5):1058–68. [JBH]

Emmons, R. A. & King, L. A. (1988) Conflict among personal strivings: Immediate and long-term implications for psychological and physical well-being. *Journal of Personality and Social Psychology* 54(6):1040–48. [JBH, CK]

Everitt, B. J. & Robbins, T. W. (2005) Neural systems of reinforcement for drug addiction: From actions to habits to compulsion. *Nature Neuroscience* 8:1481–89. [NLC]

Fabrega, H. (1997) Earliest phases in the evolution of sickness and healing. *Medical Anthropology Quarterly* 11:26–55. [SLN]

Fazio, R. H. (1986) How do attitudes guide behavior? In *The handbook of motivation and cognition, vol. 1*, ed. R. M. Sorrentino & E. T. Higgins, pp. 204–43. Guilford. [aJYH]

Fazio, R. H. (1990) Multiple processes by which attitudes guide behavior: The MODE model as an integrative framework. In: *Advances in experimental social psychology, vol. 23*, ed. M. P. Zanna, pp. 75–109. Academic Press. [aJYH]

Fedyk, M. (2014) How (not) to bring biology and psychology together. *Philosophical Studies*. doi: 10.1007/s11098-014-0297-9. Available at: http://link.springer.com/article/10.1007/s11098-014-0297-9 [ePub ahead of print] [MF]

Feldman, H. & Friston, K. J. (2010) Attention, uncertainty, and free-energy. *Frontiers in Human Neuroscience* 4:215. [GP]

Feldman, R., Gordon, I., Influs, M., Gutbir, T., & Ebstein, R. P. (2013) Parental oxytocin and early caregiving jointly shape children's oxytocin response and social reciprocity. *Neuropsychopharmacology* 38(7):1154–62. [CS]

Ferguson, M. J. (2007) On the automatic evaluation of end-states. *Journal of Personality and Social Psychology* 92:596. [BE]

Ferguson, M. J. (2008) On becoming ready to pursue a goal you don't know you have: Effects of nonconscious goals on evaluative readiness. *Journal of Personality and Social Psychology* 95:557–72. [arJYH]

Ferguson, M. J. & Bargh, J. A. (2004) Liking is for doing: The effects of goal pursuit on automatic evaluation. *Journal of Personality and Social Psychology* 87:557–72. Available at: http://www.rotman.utoronto.ca/marketing/Ferguson%20BACKGROUND%20PAPER%20-%20Liking%20Is%20For%20Doing%20-%20The%20Effects%20of%20Goal%20Pursuit.pdf. [aJYH]

Ferguson, M. J., Bargh, J. A. & Nayak, D. A. (2005) After-affects: How automatic evaluations influence the interpretation of subsequent, unrelated stimuli. *Journal of Experimental Social Psychology* 41:182–91. [aJYH]

Festinger, L. A. (1957) *A theory of cognitive dissonance*. Stanford University Press. [PC, AF]

Fishbach, A. & Converse, B. A. (2010) Identifying and battling temptation. In: *Handbook of self-regulation: Research, theory and applications*, 2nd ed., ed. K. D. Vohs & R. F. Baumeister, pp. 244–60. Guilford. [AF]

Fishbach, A. & Dhar, R. (2005) Goals as excuses or guides: The liberating effect of perceived goal progress on choice. *Journal of Consumer Research* 32:370–77. Available at: http://faculty.chicagobooth.edu/ayelet.fishbach/research/FD_JCR_05.pdf. [aJYH]

Fishbach, A., Eyal, T. & Finkelstein, S. R. (2010) How positive and negative feedback motivate goal pursuit. *Social and Personality Psychology Compass* 4:517–30. [AF]

Fishbach, A. & Ferguson, M. F. (2007) The goal construct in social psychology. In *Social psychology: Handbook of basic principles,* 2nd ed., ed. A. W. Kruglanski & E. T. Higgins, pp. 490–515. Guilford. [aJYH]

Fishbach, A., Friedman, R. S. & Kruglanski, A. W. (2003) Leading us not unto temptation: Momentary allurements elicit overriding goal temptation. *Journal of Personality and Social Psychology* 84:296–309. Available at: http://faculty.chicagobooth.edu/ayelet.fishbach/research/tg.pdf. [aJYH, GBM]

Fishbach, A. & Shah, J. Y. (2006) Self-control in action: Implicit dispositions toward goals and away from temptations. *Journal of Personality and Social Psychology* 90:820–32. Available at: http://faculty.chicagobooth.edu/ayelet.fishbach/research/FS_JPSP2006.pdf. [aJYH]

Fishbach, A. & Trope, Y. (2005) The substitutability of external control and self-control in overcoming temptation. *Journal of Experimental Social Psychology* 41:256–70. Available at: http://faculty.chicagobooth.edu/ayelet.fishbach/research/subst.pdf. [aJYH]

Fishbach, A. Zhang, Y. & Koo, M. (2009) The dynamics of self-regulation. *European Review of Social Psychology* 20:15–344. [CK]

Fiske, S. T. (2013) *Social cognition: From brains to culture*. SAGE Publications Limited. [rJYH]

Fiske, S. T., Cuddy, A. J. C. & Glick, P. (2007) Universal dimensions of social perception: Warmth and competence. *Trends in Cognitive Science* 11:77–83. [aJYH]

Fitzsimons, G. M. & Bargh, J. A. (2003) Thinking of you: Nonconscious pursuit of interpersonal goals associated with relationship partners. *Journal of Personality and Social Psychology* 84:148–64. Available at: http://www.yale.edu/acmelab/articles/Fitzsimons_Bargh_JPSP2003.pdf. [aJYH]

Fitzsimons, G. M. & Fishbach, A. (2010) Shifting closeness: Interpersonal effects of personal goal progress. *Journal of Personality and Social Psychology*. 98:535–49. Available at: http://faculty.chicagobooth.edu/ayelet.fishbach/research/FF_JPSP09.pdf. [aJYH]

Fitzsimons, G. M. & Shah, J. Y. (2008) How goal instrumentality shapes relationship evaluations. *Journal of Personality and Social Psychology*, 95:319–37. Available at: http://www.ncbi.nlm.nih.gov/pubmed/18665705. [aJYH]

Foa, E. B., Kozak, M. J., Goodman, W. K., Hollander, E., Jenike, M. A. & Rasmussen, S. A. (1995) DSM-IV field trial: Obsessive-compulsive disorder. *American Journal of Psychiatry* 152:90–96. [NLC]

Förster, J., Liberman, N. & Friedman, R. S. (2007) Seven principles of goal activation: A systematic approach to distinguishing goal priming from priming of non-goal constructs. *Personality and Social Psychology Review* 11:211–33. [aJYH, GBM]

Förster, J., Liberman, N. & Higgins, E. T. (2005) Accessibility from active and fulfilled goals. *Journal of Experimental Social Psychology* 41:220–39. Available at: http://www.columbia.akadns.net/cu/psychology/higgins/papers/forsterliberman-higgins2005.pdf. [arJYH]

Freud, S. (1901) *The psychopathology of everyday life,* trans. A. A. Brill, T. Fisher Unwin. [aJYH]

Freud, S. (1915) The unconscious. In: *Penguin Freud library. On metapsychology, vol. 11.* Penguin 1984. [TAM]

Frijda, N. H. (1986) *The emotions*. Cambridge University Press. [EB-M]

Friston, K. (2005) A theory of cortical responses. *Philosophical Transactions of the Royal Society of London* 360:815–36. [GP]

Friston, K. (2010) The free-energy principle: A unified brain theory? *Nature Reviews Neuroscience* 11:127–38. [GP]

Frith, C. D., Blakemore, S.-J. & Wolpert, D. M. (2000) Abnormalities in the awareness and control of action. *Philosophical Transactions of the Royal Society of London* 355:1771–88. [aJYH, GP]

Gable, P. & Harmon-Jones, E. (2010) The motivational dimensional model of affect: Implications for breadth of attention, memory, and cognitive categorisation. *Cognition & Emotion* 24(2):322–37. [EB-M]

Gazzaniga, M. S. (1984) *The social brain*. Basic Books. [aJYH]

Gazzaniga, M. S. (2005) *The ethical brain*. The Dana Press. [rJYH]

Gendler, T. (2008a) Alief and belief. *Journal of Philosophy* 105(10):634–63. [BH]

Gendler, T. (2008b) Alief in action (and reaction). *Mind and Language* 23 (5):552–85. [BH]

Gentner, D. & Grudin J. (1985) The evolution of mental metaphors in psychology: A 90-year retrospective. *American Psychologist* 40:181–92. [rJYH]

Gibson, J. J. (1979) *The ecological approach to visual perception*. Erlbaum. [GP]

Gilbert, R. M. (1972) Variation and selection of behavior. In: *Reinforcement: Behavioral analyses,* ed. R. M. Gilbert & J. R. Millenson, pp. 263–76. Academic Press. [GA]

Gillan, C. M., Morein-Zamir, S., Urcelay, G. P., Sule, A., Voon, V., Apergis-Schoute, A. M., Fineberg, N. A., Sahakian, B. J. & Robbins, T. W. (2013) Enhanced avoidance habits in obsessive-compulsive diesorder. *Biological Psychiatry*. doi: 10.1016/j.biopsych.2013.02.002. [ePub ahead of print] [NLC]

Glenberg, A. (1997) What memory is for. *Behavioral and Brain Sciences* 20:1–55. [GP]

Godfrey-Smith, P. (2000) The replicator in retrospect. *Biology and Philosophy* 15:403–23. [BN]

Godwin, C. A., Gazzaley, A. & Morsella, E. (2013) Homing in on the brain mechanisms linked to consciousness: Buffer of the perception-and-action interface. In: *The unity of mind, brain and world: Current perspectives on a science of consciousness*, ed. A. Pereira & D. Lehmann, pp. 43–76. Cambridge University Press. [aJYH]

Gollwitzer, P. M. (1999) Implementation intentions: Strong effects of simple plans. *American Psychologist* 54:493–503. [arJYH]

Gollwitzer, P. M. & Moskowitz, G. B. (1996) Goal effects on action and cognition. In: *Social psychology: Handbook of basic principles*, ed. E. T. Higgins & A. W. Kruglanski, pp. 361–99. Guilford. [arJYH]

Goodale, M. A. (2011) Transforming vision into action. *Vision Research* 51:1567–87. [BN]

Goodale, M. A., Milner, A. D., Jakobson, L. S. & Carey, D. P. (1991) A neurological dissociation between perceiving objects and grasping them. *Nature* 349:154–56. [aJYH]

Goodwin, B. (2001) *How the leopard changed its spots.* Princeton University Press. [CK]

Goschke, T. & Kuhl, J. (1993) Representation of intentions: Persisting activation in memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 19:1211–26. [aJYH]

Gosling, P., Denizeau, M. & Oberlé, D. (2006) Denial of responsibility: A new mode of dissonance reduction. *Journal of Personality and Social Psychology* 90:722–33. [PC]

Graybiel, A. M. & Rauch, S. L. (2000) Toward a neurobiology of obsessive-compulsive disorder. *Neuron* 28:343–47. [NLC]

Greenberg, J., Koole, S. L. & Pyszczynski, T. eds. (2004) *Handbook of experimental existential psychology.* Guilford. [ACK]

Greene, J. D., Nystrom, L. E., Engell, A. D., Darley, J. M. & Cohen, J. D. (2004) The neural bases of cognitive conflict and control in moral judgment. *Neuron* 44(2):389–400. Available at: http://www.ncbi.nlm.nih.gov/pubmed/15473975. [CS]

Greenwald, A. G. (2004) The resting parrot, the dessert stomach, and other perfectly defensible theories. In: *The yin and yang of social cognition: Perspectives on the social psychology of thought systems,* ed. J. T. Jost, M. R. Banaji & D. A. Prentice, pp. 275–85. American Psychological Association. [ACK]

Greenwald, A. G. & Banaji, M. R. (1995) Implicit social cognition. *Psychological Review* 102:4–27. [aJYH]

Griskevicius, V., Tybur, J. M., Gangestad, S. W., Perea, E. F., Shapiro, J. R. & Kenrick, D. T. (2009) Aggress to impress: Hostility as an evolved context-dependent strategy. *Journal of Personality and Social Psychology* 96:980–94. Available at: http://www.csom.umn.edu/assets/125344.pdf. [arJYH]

Gross, J. J. (2002) Emotion regulation: Affective, cognitive, and social consequences. *Psychophysiology* 39(3):281–91. Available at: http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=12212647. [CS]

Grush, R. (2004) The emulation theory of representation: Motor control, imagery, and perception. *Behavioral and Brain Sciences* 27(3):377–96. [BH]

Haber, S. N., Fudge, J. L. & McFarland, N. R. (2000) Striatonigrostriatal pathways in primates form an ascending spiral from the shell to the dorsolateral striatum. *Journal of Neuroscience* 20(6):2369–82. [CPM]

Haggard, P. & Clark, S. (2003) Intentional action: Conscious experience and neural prediction. *Consciousness and Cognition* 12:695–707. [TAM]

Hagger, M. S., Wood, C., Stiff, C & Chatzisarantis, N. L. (2010) Ego depletion and the strength model of self-control: A meta-analysis. *Psychological Bulletin* 4:495–525. [rJYH]

Hamilton, D. L., Katz, L. B. & Leirer, V. O. (1980) Cognitive representation of personality impressions: Organizational processes in first impression formation. *Journal of Personality and Social Psychology* 39:1050–63. [aJYH]

Hamilton, W. D. (1964) The genetical evolution of social behaviour. I & II. *Journal of Theoretical Biology* 7(1):1–52. [DC-B, BM]

Han, S. D., Drake, A. I., Cessante, L. M., Jak, A. J., Houston, W. S., Delis, D. C., Filoteo, J. V. & Bondi, M. W. (2007) Apolipoprotein E and traumatic brain injury in a military population: Evidence of a neuropsychological compensatory mechanism? *Journal of Neurology, Neurosurgery, and Psychiatry* 78:1103–108. [WvH]

Hannan, M. T. & Freeman, J. (1977) The population ecology of organization. *American Journal of Sociology* 82:929–64. [rJYH]

Harber, K. D., Yeung, D. & Iacovelli, A. (2011) Psychosocial resources, threat, and the perception of distance and height: Support for the resources and perception model. *Emotion* 11:1080–90. [GBM]

Hassin, R. R. (2005) Nonconscious control and implicit working memory. In: *The new unconscious*, ed. R. R. Hassin, J. S. Uleman & J. A. Bargh, pp. 196–222. Oxford University Press. [aJYH]

Hastie, R. & Park, B. (1986) The relationship between memory and judgment depends on whether the judgment task is memory-based or online. *Psychological Review* 93:258–68. [aJYH]

Hastorf, A. H. & Cantril, H. (1954) They saw a game: A case study. *Journal of Abnormal and Social Psychology* 49:129–34. [aJYH]

Heckhausen, H. (1991) *Motivation and action*. Springer. [aJYH]

Heckhausen, H. & Gollwitzer, P. M. (1987) Thought contents and cognitive functioning in motivational versus volitional states of mind. *Motivation and Emotion* 11(2):101–20. [EB-M]

Helzer, E. G. & Pizarro, D. A. (2011) Dirty liberals!: Reminders of physical cleanliness influence moral and political attitudes. *Psychological Science* 22:517–22. [SLN]

Higgins, E. T. (1987) Self-discrepancy: A theory relating self and affect. *Psychological Review* 94:319–40. [AF]

Higgins, E. T. (1996) Knowledge activation: Accessibility, applicability, and salience. In: *Social psychology: Handbook of basic principles,* ed. E. T. Higgins & A. W. Kruglanski, pp. 133–68. Guilford. [aJYH]

Higgins, E. T. (1997) Beyond pleasure and pain. *The American Psychologist* 52 (12):1280–300. [EB-M]

Higgins, E. T. (2012) *Beyond pleasure and pain: How motivation works.* Oxford University Press. [BE]

Hill, S. E. & Durante, K. M. (2011) Courtship, competition, and the pursuit of attractiveness: Mating goals facilitate health-related risk taking and strategic risk suppression in women. *Personality and Social Psychology Bulletin* 37:383–94. [aJYH]

Hirsh, J. B. (2010) The weight of being: Psychological perspectives on the existential moment. *New Ideas in Psychology* 28:28–36. [JBH]

Hirsh, J. B., Mar, R. A. & Peterson, J. B. (2012) Psychological entropy: A framework for understanding uncertainty-related anxiety. *Psychological Review* 119 (2):304–20. [JBH]

Hoffman, R. R., Cochran, E. L. & Nead, J. M. (1990) Cognitive metaphors in experimental psychology. In: *Metaphors in the History of Psychology,* ed. D. E. Leary, pp. 173–229. Cambridge University. [rJYH]

Hoffree, G. & Winkielman, P. (2012) On (not) knowing and feeling what we want and like. In: *Handbook of self-knowledge,* ed. S. Vazire & T. D. Wilson, pp. 210–24. Guilford Publication. [JBH]

Hofmann, W., Friese, M. & Strack, F. (2009) Impulse and self-control from a dual-systems perspective. *Perspectives on Psychological Science* 4(2):162–76. [CS]

Holland, R. W., Hendriks, M. & Aarts, H. (2005) Smells like clean spirit: Nonconscious effects of scent on cognition and behavior. *Psychological Science* 16:689–93. Available at: http://goallab.nl/publications/documents/Holland,%20Hendriks,%20Aarts%20%282005%29%20-%20noncsious%20effects%20of%20scent%20on%20behavior.pdf. [aJYH]

Hommel, B., Musseler, J., Aschersleben, G. & Prinz, W. (2001) The theory of event coding (TEC): A framework for perception and action planning. *Behavioral and Brain Sciences* 24(5):849–78. [GP]

Hordern, P. J. (1972) Religious conceptions and the world of nature in ancient Egypt. Available at: http://digitalcommons.mcmaster.ca/opendissertations/7701/. [TAM]

Huang, S. (2012) The molecular and mathematical basis of Waddington's epigenetic landscape: A framework for post-Darwinian biology? *Bioessays* 34:149–57. [TAM]

Huebner, B. (2013) *Macrocognition*. Oxford University Press. [BH]

Hume, D. (1739) A treatise of human nature. Available at: http://www.gutenberg.org/files/4705/4705-h/4705-h.htm. [DVB]

Ismael, J. (2007) *The situated self.* Oxford University Press. [NW]

Ismael, J. (2011) Self-organization and self-governance. *Philosophy of the Social Sciences* 41(3):327–51. [NW]

Jablonka, E. & Lamb, M. J. (2005) *Evolution in four dimensions: Genetic, epigenetic, behavioral, and symbolic variation in the history of life*. MIT Press. [MF]

Jacob, F. (1977) Evolution and tinkering. *Science* 196:1161–66. Available at: http://www.sciencemag.org/content/196/4295/1161.full.pdf. [aJYH]

James, W. (1890) *Principles of psychology, 2 vols*. Henry Holt. [TAM]

James, W. (1890/1981) *Principles of psychology,* Harvard University Press. [aJYH]

Jeannerod, M. (1997) *The cognitive neuroscience of action.* Blackwell. [BN]

Jencks, C. (1990) Varieties of altruism. In: *Beyond self-interest*, ed. J. J. Mansbridge, pp. 53–67. University of Chicago Press. [aJYH]

Job, V., Dweck, C. S. & Walton, G. M. (2010) Ego depletion – Is it all in your head?: Implicit theories about willpower affect self-regulation. *Psychological Science* 21:1686–93. [rJYH]

Johnson, E. J. & Goldstein, D. (2003) Do defaults save lives? *Science* 302:1338–39. [rJYH]

Jost, J. T., Kay, A. C. & Thorisdottir, H. eds. (2009) *Social and psychological bases of ideology and system justification.* Oxford University Press. [ACK]

Jost, J. T., Pelham, B. W., Sheldon, O. & Sullivan, B. N. (2003) Social inequality and the reduction of ideological dissonance on behalf of the system: Evidence of enhanced system justification among the disadvantaged. *European Journal of Social Psychology* 33:13–36. [rJYH]

Jost, J. T., Pietrzak, J., Liviatan, I., Mandisodza, A. N. & Napier, J. L. (2008) System justification as conscious and nonconscious goal pursuit. In: *Handbook of motivation science*, ed. J. Shah & W. Gardner, pp. 591–605. Guilford. [aJYH, ACK]

Jost, J. T. & Thompson, E. P. (2000) Group-based dominance and opposition to equality as independent predictors of self-esteem, ethnocentrism, and social policy attitudes among African Americans and European Americans. *Journal of Experimental Social Psychology* 36:209–32. [ACK]

Jung, C. G. (1939) *The integration of the personality.* (trans. S. Dell). Farrar & Rinehart. [JBH]

Kable, J. W. & Glimcher, P. W. (2007) The neural correlates of subjective value during intertemporal choice. *Nature Neuroscience* 10(12):1625–33. [GA, DS]

Kay, A., Jimenez, M. C. & Jost, J. T. (2002) Sour grapes, sweet lemons, and the anticipatory rationalization of the status quo. *Personality and Social Psychology Bulletin* 28:1300–12. [ACK]

Kay, A. C., Gaucher, D., Peach, J. M., Friesen, J., Laurin, K., Zanna, M. P. & Spencer, S. J. (2009) Inequality, discrimination, and the power of the status quo: Direct evidence for a motivation to view what is as what should be. *Journal of Personality and Social Psychology* 97:421–34. [ACK]

Kay, A. C., Wheeler, S. C., Bargh, J. A. & Ross, L. (2004) Material priming: The influence of mundane physical objects on situational construal and competitive behavioral choice. *Organizational Behavior and Human Decision Processes* 95:83–96. Available at: http://www.sciencedirect.com/science/article/B6WP2-CVX3K0-/2/e5b3aebe6709c76f8834518e97430d6e. [aJYH]

Keltner, D. & Gross, J. J. (1999) Functional accounts of emotions. *Cognition & Emotion* 13(5):467–80. [EB-M]

Kennedy, D., ed. (2005) 125 questions: What don't we know? [Special issue]. *Science* 309:5731. [aJYH]

Kenrick, D. T., Griskevicius, V., Neuberg, S. L. & Schaller, M. (2010) Renovating the pyramid of needs: Contemporary extensions built upon ancient foundations. *Perspectives on Psychological Science* 5:292–314. Available at: http://www.csom.umn.edu/assets/144040.pdf. [arJYH, DVB, SLN]

Kihlstrom, J. F. (1987) The cognitive unconscious. *Science* 237:1445–52. [TAM]

Killeen, P. R. (2001) The four causes of behavior. *Current Directions in Psychological Science* 10:36–40. Available at: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2600470/. [aJYH]

Kim, P., Feldman, R., Mayes, L. C., Eicher, V., Thompson, N., Leckman, J. F. & Swain, J. E. (2011) Breastfeeding, brain activation to own infant cry, and maternal sensitivity. *Journal of Child Psychology and Psychiatry* 52(8):907–915. Available at: http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=21501165. [CS]

Kirkwood, T. B. L. (1977) Evolution of aging. *Nature* 270:301–304. [BM]

Kirkwood, T. B. L. & Rose, M. R. (1991) Evolution of senescence: Late survival sacrificed for reproduction. *Philosophical Transactions of the Royal Society, London Series* 332:15–24. [BM]

Kleiman, T. & Hassin, R. R. (2011) Non-conscious goal conflicts. *Journal of Experimental Social Psychology* 47:521–32. [aJYH]

Knowlton, B. J., Mangels, J. A. & Squire, L. R. (1996) A neostriatal habit learning system in humans. *Science* 273:1399–402. [CPM]

Kober, H., Kross, E. F., Mischel, W., Hart, C. L. & Ochsner, K. N. (2010) Regulation of craving by cognitive strategies in cigarette smokers. *Drug and Alcohol Dependence* 106:52–55. [CK]

Koechlin, E., Basso, G., Pietrini, P., Panzer, S. & Grafman, J. (1999) The role of the anterior prefrontal cortex in human cognition. *Nature* 399(6732):148–51. [JBH]

Kolb, B., Gorny, G., Li, Y. L., Samaha, A. N. & Robinson, T. E. (2003) Amphetamine or cocaine limits the ability of later experience to promote structural plasticity in the neocortex and nucleus accumbens. *Proceedings of the National Academy of Sciences of the United States of America* 100:10523–28. [CPM]

Köpetz, C., Faber, T., Fishbach, A. & Kruglanski, A. W. (2011) The multifinality constraints effect: How goal multiplicity narrows the means set to a focal end. *Journal of Personality and Social Psychology* 100:810–26. [AF, CK]

Kruglanski, A. W., Kopetz, C., Belanger, J., Chun, W. Y., Orehek, E. & Fishbach, A. (2013) Features of multifinality: Effects of goal plurality on means preferences. *Personality and Social Psychology Review* 17:22–39. [CK]

Kruglanski, A. W., Shah, J. Y., Fishbach, A., Friedman, R., Chun, W. Y. & Sleeth-Keppler, D. (2002) A theory of goal systems. In: *Advances in experimental social psychology, vol. 34*, ed. M. P. Zanna, pp. 331–78. Academic Press. Available at: http://faculty.chicagobooth.edu/ayelet.fishbach/research/advances.pdf. [arJYH, AF, JBH, CK, GBM]

Kuhl, J. (1984) Volitional aspects of achievement motivation and learned helplessness: Toward a comprehensive theory of action-control. In: *Progress in experimental personality research, vol. 13,* ed. B. A. Maher, pp. 99–171. Academic Press. [aJYH]

Kuhl, J. & Weiss, M. (1994) Performance deficits following uncontrollable failure: Impaired action control or global attributions and generalized expectancy deficits? In *Volition and personality: Action versus state orientation,* ed. J. Kuhl & J. Beckmann, pp. 317–28. Hogrefe & Huber. [aJYH]

Kühn, S., Brass, M. & Haggard, P. (2013) Feeling in control: Neural correlates of experience of agency. *Cortex* 49(7):1935–42. doi: 10.1016/j.cortex.2012.09.002. [TAM]

Kunda, Z. (1990) The case for motivated reasoning. *Psychological Bulletin* 108:480–98. Available at: http://www.psych.utoronto.ca/users/peterson/psy430s2001/Kunda%20Z%20Motivated%20Reasoning%20Psych%20Bull%201990.pdf. [aJYH]

Kunda, Z. & Spencer, S. J. (2003) When do stereotypes come to mind and when do they color judgment? A goal-based theoretical framework for stereotype activation and application. *Psychological Bulletin* 129:522–44. Available at: http://cogsci.uwaterloo.ca/ziva/psychbul2003.pdf. [aJYH]

Kunda, Z. & Thagard, P. (1996) Forming impressions from stereotypes, traits, and behaviors: A parallel-constraint-satisfaction theory. *Psychological Review* 103(2):284–308. [rJYH]

Kurzban, R. (2012) *Why everyone (else) is a hypocrite: Evolution and the modular mind*. Princeton University Press. [DC-B]

Kurzban, R., Dukes, A. & Weeden, J. (2010) Sex, drugs and moral goals: Reproductive strategies and views about recreational drugs. *Proceedings of the Royal Society B: Biological Sciences* 277(1699):3501–508. [DC-B]

Kushnir, T. (2012) Developing a concept of choice. In: *Advances in child development and behavior: Rational constructivism in cognitive development,* ed. F. Xu & T. Kushnir, pp. 193–213. Academic Press. [MF]

Lashley, K. S. (1942) The problem of cerebral organization in vision. In: *Visual mechanisms. Biological symposia, vol. 7,* ed. H. Kluver, pp. 301–22. Cattell Press. [aJYH]

Laurin, K., Kay, A. C. & Fitzsimons, G. J. (2012) Reactance versus rationalization: Divergent responses to Constrained Freedom. *Psychological Science* 23:205–209. [ACK]

Ledgerwood, A., Mandisodza, A., Jost, J. T. & Pohl, M. (2011) Working for the system: Motivated defense of meritocratic beliefs. *Social Cognition* 29:322–40. [ACK]

Levy, D. J. & Glimcher, P. W. (2012) The root of all value: A neural common currency for choice. *Current Opinion in Neurobiology* 22:1027–38. [DS]

Levy, N. (2012) Consciousness, implicit attitudes, and moral responsibility. *Noûs*. doi: 10.1111/j.1468-0068.2011.00853.x. [NW]

Lewin, K. (1926) Vorsatz, wille, und bedürfnis [Intention, will, and need]. *Psychologische Forschung* 7:330–85. [aJYH]

Lewin, K. (1935) *A dynamic theory of personality*. McGraw-Hill. [aJYH]

Lhermitte, F. (1986) Human anatomy and the frontal lobes: Part II: Patient behavior in complex and social situations: The "environmental dependency syndrome." *Annals of Neurology* 19:335–43. [aJYH]

Liviatan, I. & Jost, J. T. (2014) A social-cognitive analysis of system justification goal striving. *Social Cognition* 32(2):95–129 [ACK]

Locke, E. A. & Latham, G. P. (1990) *A theory of goal setting and task performance*. Prentice-Hall. [aJYH]

Locke, E. A. & Latham, G. P. (2002) Building a practically useful theory of goal setting and task performance: A 35 year odyssey. *American Psychologist* 57:705–17. [aJYH]

Loersch, C. & Payne, B. K. (2012) On mental contamination: The role of (mis) attribution in behavior priming. *Social Cognition* 30:241–52. [aJYH]

Loewenstein, G. (1996) Out of control: Visceral influences on behavior. *Organizational Behavior and Human Decision Processes* 65:272–92. [aJYH, NLC]

Loftus, E. F. & Klinger, M. R. (1992) Is the unconscious smart or dumb? *American Psychologist* 47:761–65. Available at: http://econshp2.econ.kuleuven.ac.be/consumerscience/SiteFiles/loftus-klinger-AmPsy-1992.pdf. [aJYH]

Longo, G. & Tendero, P.-E. (2007) The differential method and the causal incompleteness of programming theory in molecular biology. *Foundations of Science* 12:337–66. [TAM]

Lord, C. G., Ross, L. & Lepper, M. R. (1979) Biased assimilation and attitude polarization: The effects of prior theories on subsequently considered evidence. *Journal of Personality and Social Psychology* 37:2098–109. Available at: http://www.psych.umn.edu/courses/spring07/borgidae/psy5202/readings/lord,%20ross%20&%20lepper%20%281979%29.pdf. [aJYH]

Mack, A. (2003) Inattentional blindness: Looking without seeing. *Current Directions in Psychological Science* 12:180–84. Available at: http://www.unc.edu/~pcg/225/documents/InattentionalBLindness2003.pdf. [aJYH]

Mackie, D. M. & Hamilton, D. L. (1993) *Affect, cognition, and stereotyping: Interactive processes in group perception.* Academic Press. [EB-M]

Macphail, E. (1998) *The evolution of consciousness*. Oxford University Press. [aJYH]

Macrae, C. N., Bodenhausen, G. V., Milne, A. B., Thorn, T. M. J. & Castelli, L. (1997) On the activation of social stereotypes: The moderating role of processing objectives. *Journal of Experimental Social Psychology* 33:471–89. Available at: http://www.sciencedirect.com/science?_ob=ArticleURL&_udi=B6WJB-45KV10P-8&_user=483702&_coverDate=09%2F30%2F1997&_rdoc=1&_fmt=high&_orig=search&_origin=search&_sort=d&_docanchor=&view=c&_searchStrId=1577489557&_rerunOrigin=google&_acct=C000022720&_version=1&_urlVersion=0&_userid=483702&md5=9afd6ad0e00045664577fae40a65ceda&searchtype=a. [aJYH]

Maddux, W. M., Barden, J., Brewer, M. B. & Petty, R. E. (2005) Saying no to negativity: The effects of context and motivation to control prejudice on automatic evaluative responses. *Journal of Experimental Social Psychology* 41:19–35. Available at: http://faculty.insead.edu/maddux/personal/documents/JESP-MotivePaper.pdf. [aJYH]

Maner, J. K., DeWall, C. N. & Gailliot, M. T. (2008) Selective attention to signs of success: Social dominance and early stage interpersonal perception. *Personality and Social Psychology Bulletin* 34:488–501. Available at: http://www.psy.fsu.edu/faculty/maner/maner%20dewall%20gailliot%20PSPB%20attention%20dominance.doc. [aJYH]

Maner, J. K., Gailliot, M. T. & DeWall, C. N. (2007) Adaptive attentional attunement: Evidence for mating-related perceptual bias. *Evolution and Human Behavior* 28:28–36. [GBM]

Maner, J. K., Kenrick, D. T., Becker, D. V., Robertson, T. E., Hofer, F., Neuberg, S. L., Delton, A. W., Butner, J. & Schaller, M. (2005) Functional projection: How fundamental social motives can bias interpersonal perception. *Journal of Personality and Social Psychology* 88:63–78. [SLN]

Mansbridge, J. J. (1990) The rise and fall of self-interest in the explanation of political life. In: *Beyond self-interest*, ed. J. J. Mansbridge, pp. 3–22. University of Chicago Press. [aJYH]

Marchetti, C. & Della Sala, S. D. (1998) Disentangling the alien and anarchic hand. *Cognitive Neuropsychiatry* 3:191–207. [aJYH]

Marien, H., Custers, R., Hassin, R. R. & Aarts, H. (2012) Unconscious goal activation and the hijacking of the executive function. *Journal of Personality and Social Psychology* 103:399–415. [aJYH]

Marsh, R. L., Hicks, J. L. & Bink, M. L. (1998) Activation of completed, uncompleted, and partially completed intentions. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 24:350–61. Available at: http://dtserv2.compsy.uni-jena.de/ss2008/psychnf_uj/72076412/content.nsf/Pages/88222F318051357DC1257441003DEEA8/$FILE/marsh%20et%20al%201998%20completed%20and%20uncompleted%20intentions.pdf. [aJYH]

Martin, L. L. & Tesser, A. (2009) Five markers of motivated behavior. In: *The psychology of goals*, ed. G. B. Moskowitz and H. Grant, pp. 257–76. Guilford. [GBM]

Martin, L. L., Ward, D. W., Achee, J. W. & Wyer, R. S. (1993) Mood as input: People have to interpret the motivational implications of their moods. *Journal of Personality and Social Psychology* 64:317–26. [EB-M]

Martindale, C. (1980) Subselves. In: *Review of personality and social psychology*, ed. L. Wheeler, pp. 193–218. Sage. [DVB]

Maslow, A. H. (1970) *Motivation and personality*. Harper & Row. [JBH]

Mayes, L. C., Swain, J. E. & Leckman, J. F. (2005) Parental attachment systems: Neural circuits, genes, and experiential contributions to parental engagement. *Clinical Neuroscience Research* 4(5–6):301–13. [CS]

Maynard-Smith, J. (1982) *Evolution and the theory of games*. Cambridge University Press. [DS]

Mayr, E. (1976) *Evolution and the diversity of life*. Harvard University Press. [aJYH]

Mazzone, M. & Campisi, E. (2013) Distributed intentionality. A model of intentional behavior in humans. *Philosophical Psychology* 26: 267–90. [MM]

McCoy, S. T. & Major, B. (2007) Priming meritocracy and the psychological justification of inequality. *Journal of Experimental Social Psychology* 43:341–51. [ACK]

McCulloch, K. D., Ferguson, M. J., Kawada, C. & Bargh, J. A. (2008) Taking a closer look: On the operation of nonconscious impression formation. *Journal of Experimental Social Psychology* 44:614–23. Available at: http://www.yale.edu/acmelab/articles/McCulloch_et_al_JESP.pdf. [aJYH]

McGregor, I. & Little, B. R. (1998) Personal projects, happiness, and meaning: On doing well and being yourself. *Journal of Personality and Social Psychology* 74 (2):494–512. [JBH]

McKay, R. & Dennett, D. (2009) The evolution of misbelief: Our evolving beliefs about evolved misbelief. *Behavioral and Brain Sciences* 32:493–561. Available at: http://homepage.mac.com/ryantmckay/McKay&Dennett09.pdf. [aJYH]

Medawar, P. B. (1952) *An unsolved problem of biology*. H. K. Lewis. [WvH]

Memelink, J. & Hommel, B. (2013) Intentional weighting: A basic principle in cognitive control. *Psychology Research* 77:249–59. [GP]

Mercier, H. & Sperber, D. (2011) Why do humans reason? Arguments for an argumentative theory. *Behavioral and Brain Sciences* 34:57–74. [aJYH]

Merker, B. (2007) Consciousness without a cerebral cortex: A challenge for neuroscience and medicine. *Behavioral and Brain Sciences* 30:63–134. [aJYH]

Milgram, S. (1963) Behavioral study of obedience. *Journal of Abnormal and Social Psychology* 67:371–78. [aJYH]

Miller, D. T. (1999) The norm of self-interest. *American Psychologist* 54:1–8. Available at: http://cepr.org/meets/wkcn/3/3509/papers/Miller.pdf. [aJYH]

Miller, E. K. & Cohen, J. D. (2001) An integrative theory of prefrontal cortex function. *Annual Review of Neuroscience* 24:167–202. [GP]

Miller, G. A., Galanter, E. & Pribram, K. H. (1960) *Plans and the structure of behavior*. Holt, Rinehart and Winston. [GP]

Miller, S. L., Maner, J. K. & Becker, D. V. (2010) Self-protective biases in group categorization: Threat cues shape the psychological boundary between "us" and "them." *Journal of Personality and Social Psychology* 99:62–77. [SLN]

Millikan, R. (1995) Pushmi-pullyu representations. *Philosophical Perspectives* 9:185–200. [BH]

Milner, A. D. & Goodale, M. A. (1995) *The visual brain in action*. Oxford University Press. [aJYH]

Minsky, M. (1986) *The society of mind*. Simon and Schuster. [DVB]

Mirouze, M. (2012) The small RNA-based odyssey of epigenetic information in plants: From cells to species. *DNA Cell Biology* 31:1650–56. doi: 10.1089/dna.2012.1681. [TAM]

Mischel, W. (1973) Toward a cognitive social learning reconceptualization of personality. *Psychological Review* 80:252–83. [aJYH]

Mischel, W. (1997) Was the cognitive revolution just a detour on the road to behaviorism? On the need to reconcile situational control and personal control. In: *Advances in social cognition, vol. 10,* ed. R. S. Wyer, Jr., pp. 181–86. Erlbaum. [GA]

Mischel, W. (2004) Toward an integrative science of the person. *Annual Review of Psychology* 55:1–22. [RFB]

Mischel, W. & Shoda, Y. (1995) A cognitive-affective system theory of personality: Reconceptualizing situations, dispositions, dynamics, and invariance in personality structure. *Psychological Review* 102:246–68. [aJYH]

Mischel, W., Shoda, Y. & Rodriguez, M. L. (1989) Delay of gratification in children. *Science* 244(4907):933–38. [AF]

Molden, D. C., Hui, C. M., Scholar, A. A., Meier, B. P., Noreen, E. E., D'Agostino, P. R. & Maritn, V. (2012) Motivational versus metabolic effects of carbohydrates on self-control. *Psychological Science* 23(10):1137–44. [rJYH]

Monin, B. & Miller, D. T. (2001) Moral credentials and the expression of prejudice. *Journal of Personality and Social Psychology* 81:33–43. Available at: http://www-psych.stanford.edu/~monin/. [aJYH, AF]

Montague, P. R. & King-Casas, B. (2007) Efficient statistics, common currencies and the problem of reward-harvesting. *Trends in Cognitive Sciences* 11:514–19. [GP]

Moreno, K. N. & Bodenhausen, G. V. (2001) Intergroup affect and social judgment: Feelings as inadmissible information. *Group Processes & Intergroup Relations* 4 (1):21–29. [EB-M]

Morewedge, C. K. & Kahneman, D. (2010) Associative processes in intuitive judgment. *Trends in Cognitive Sciences* 14:435–40. Available at: http://www.andrew.cmu.edu/user/morewedg/personal/papers/IntuitiveJudgment.pdf. [aJYH]

Morgan, D., Grant, K. A., Gage, H. D., Mach, R. H., Kaplan, J. R., Prioleau, O., Nader, S. H., Buchheimer, N., Ehrenkaufer, R. L. & Nader, M. A. (2002) Social dominance in monkeys: Dopamine D2 receptors and cocaine self-administration. *Nature Neuroscience* 5:169–74. [WvH]

Morin, A. & Michaud, J. (2007) Self-awareness and the left inferior frontal gyrus: Inner speech use during self-related processing. *Brain Research Bulletin* 74:387–96. [TAM]

Morsella, E. (2005) The function of phenomenal states: Supramodular interaction theory. *Psychological Review* 112:1000–21. Available at: http://www.yale.edu/acmelab/articles/Morsella_2005.pdf. [aJYH]

Morsella, E. & Bargh, J. A. (2011) Unconscious action tendencies: Sources of "un-integrated' action. In: *Handbook of social neuroscience*, ed. J. Decety & J. T. Cacioppo, pp. 335–47. Oxford University Press. Available at: http://www.yale.edu/acmelab/articles/MorsellaBargh_HandbookSocNeuro.pdf. [aJYH]

Morsella, E., Krieger, S. C. & Bargh, J. A. (2009) Minimal neuroanatomy for a conscious brain: Homing in on the networks constituting consciousness. *Neural Networks* 23:14–15. [aJYH]

Moskowitz, G. B. (2002) Preconscious effects of temporary goals on attention. *Journal of Experimental Social Psychology* 38:397–404. [GBM]

Moskowitz, G. B. & Li, P. (2011) Egalitarian goals trigger stereotype inhibition: A proactive form of stereotype control. *Journal of Experimental Social Psychology* 47(1):103–16. [GBM]

Moskowitz, G. B., Li, P., Ignarri, C. & Stone, J. (2011) Compensatory cognition associated with egalitarian goals. *Journal of Experimental Social Psychology* 47(2):365–70. [GBM]

Moskowitz, G. B., Li, P. & Kirk, E. (2004) The implicit volition model: On the preconscious regulation of temporarily adopted goals. In: *Advances in* *experimental social psychology, vol.36*, ed. M. Zanna, pp. 317–413. Academic Press. [GBM]

Moskowitz, G. B., Gollwitzer, P. M., Wasel, W. & Schaal, B. (1999) Preconscious control of stereotype activation through chronic egalitarian goals. *Journal of Personality and Social Psychology* 77:167–84. Available at: http://www.psych.nyu.edu/gollwitzer/99MosGollWasSchaal_Precons.pdf. [aJYH]

Müller, C. P. (2013) Episodic memories and their relevance for psychoactive drug use and addiction. *Frontiers in Behavioral Neuroscience* 7(34):1–13. [CPM]

Müller, C. P., Pum, M. E., Schumann, G. & Huston, J. P. (2010) The role of serotonin in drug addiction. In: *The behavioral neurobiology of serotonin*, ed. C. P. Müller & B. L. Jacobs, pp. 507–46. Academic Press. [CPM]

Müller, C. P. & Schumann, G. (2011) Drugs as instruments: A new framework for non-addictive psychoactive drug use. *Behavioural Brain Research* 34 (6):293–347. [CPM]

Murray, D. R. & Schaller, M. (2012) Threat(s) and conformity deconstructed: Perceived threat of infectious disease and its implications for conformist attitudes and behavior. *European Journal of Social Psychology* 42:180–88. [SLN]

Murray, S. L., Derrick, J. L., Leder, S. & Holmes, J. G. (2008) Balancing connectedness and self-protection goals in close relationships: A levels-of-processing perspective on risk regulation. *Journal of Personality and Social Psychology* 94:429–59. Available at: http://www.ncbi.nlm.nih.gov/pubmed/18284291. [aJYH]

Mysore, S. P. & Knudsen, E. I. (2011) The role of a midbrain network in competitive stimulus selection. *Current Opinion in Neurobiology* 21:653–60. [GP]

Nahmias, E. (2010) Scientific challenges to free will. In: *A companion to the philosophy of action*, ed. T. O'Connor & C. Sandis, pp. 345–56. Wiley-Blackwell. [NW]

Nanay, B. (2002) The return of the replicator: What is philosophically significant in a general account of replication and selection? *Biology and Philosophy* 17:109–21. [BN]

Nanay, B. (2011) Replication without replicators. *Synthese* 179:455–77. [BN]

Nanay, B. (2013) *Between perception and action*. Oxford University Press. [BN]

Nash, K., McGregor, I. & Prentice, M. (2011) Threat and defense as goal regulation: From implicit goal conflict to anxious uncertainty, reactive approach motivation, and ideological extremism. *Journal of Personality and Social Psychology* 101 (6):1291–301. [JBH]

Neisser, U. (1967) *Cognitive psychology*. Prentice-Hall. [aJYH]

Nesse, R. M. & Berridge, K. C. (1997) Psychoactive drug use in evolutionary perspective. *Science* 278:63–66. [NLC]

Nesse, R. M. & Williams, G. C. (1995) *Why we get sick: The new science of Darwinian medicine*. Times Books. [BM]

Neuberg, S. L., Becker, D. V. & Kenrick, D. T. (2013) Evolutionary social cognition. In: *Oxford handbook of social cognition.* ed. D. Carlston, pp. 656–79. Oxford University Press. [SLN]

Neuberg, S. L., Kenrick, D. T., Maner, J. K. & Schaller, M. (2004) From evolved motives to everyday mentation: Evolution, goals, and cognition. In: *Social motivation: Conscious and unconscious processes*, ed. J. Forgas & K. Williams, pp. 133–52. Cambridge University Press. [arJYH]

Neuberg, S. L. & Schaller, M. (2014) Evolutionary social cognition. In: *APA handbook of personality and social psychology, vol. 1,* ed. E. Borgida & J. A. Bargh, pp. 3–45. American Psychological Association. [SLN]

Newell, B. R. & Shanks, D. R. (2014) Unconscious influences on decision making: A critical review. *Behavioral and Brain Sciences* 37(1):1–61. [TAM, rJYH]

Niedenthal, P. M. (1990) Implicit perception of affective information. *Journal of Experimental Social Psychology* 26:505–27. [aJYH]

Nisbett, R. E. & Wilson, T. D. (1977) Telling more than we can know: Verbal reports on mental processes. *Psychological Review* 84:231–59. [aJYH]

Noble, D. (2011) Neo-Darwinism, the modern synthesis and selfish genes: Are they of use in physiology? *Journal of Physiology* 589:1007–15. [TAM]

O'Brien, L. T., Major, B. N. & O'Gilbert, P. N. (2012) Gender differences in entitlement: The roles of system justifying beliefs. *Basic and Applied Social Psychology* 34:136–45. [ACK]

Ochsner, K. N. & Gross, J. J. (2005) The cognitive control of emotion. *Trends in Cognitive Sciences* 9(5):242–49. [CS]

Ochsner, K. N., Silvers, J. A. & Buhle, J. T. (2012) Functional imaging studies of emotion regulation: A synthetic review and evolving model of the cognitive control of emotion. *Annals of the New York Academy of Science* 1251: E1–24. Available at: http://www.ncbi.nlm.nih.gov/pubmed/23025352. [CS]

Over, H. & Carpenter, M. (2009) Priming third-party ostracism increases affiliative imitation in children. *Developmental Science* 12:F1–8. [aJYH]

Panksepp, J. (1998) *Affective neuroscience. The foundations of human and animal emotions*. Oxford University Press. [aJYH]

Payne, B. K., Cheng, C. M., Govorun, O. & Stewart, B. D. (2005) An inkblot for attitudes: Affect misattribution as implicit measurement. *Journal of Personality and Social Psychology* 89(3):277–93. [EB-M]

Payne, B. K., McClernon, F. J. & Dobbins, I. G. (2007) Automatic affective responses to smoking cues. *Experimental and Clinical Psychopharmacology* 15 (4):400–409. [EB-M]

Pearson, B. J. & Sánchez, A. (2008) Regeneration, stem cells, and the evolution of tumor suppression. *Cold Spring Harbor Symposia on Quantitative Biology* 73:565–72. doi: 10.1101/sqb.2008.73.045. [TAM]

Pelloux, Y., Dilleen, R., Economidou, D., Theobald, D. & Everitt, B. J. (2012) Reduced forebrain serotonin transmission is causally involved in the development of compulsive cocaine seeking in rats. *Neuropsychopharmcology* 37:2505–14. [CPM]

Penfield, W. & Jasper, H. (1954) *Epilepsy and the functional anatomy of the human brain.* Little, Brown & Co. [aJYH]

Pennartz, C. M. (2009) Identification and integration of sensory modalities: Neural basis and relation to consciousness. *Consciousness and Cognition* 18:718–39. [TAM]

Pennisi, E. (2012) Genomics. ENCODE project writes eulogy for junk DNA. *Science* 337:1159–61. [TAM]

Pervin, L. A. (1989, Ed.) *Goal concepts in personality and social psychology*. Erlbaum. [aJYH]

Pessiglione, M., Schmidt, L., Draganski, B., Kalisch, R., Lau, H., Dolan, R. & Frith, C. (2007, 12 April) How the brain translates money into force: A neuroimaging study of subliminal motivation. *Science* 316:904–906. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2631941. [aJYH]

Peterson, J. B. (1999) *Maps of meaning: The architecture of belief*. Routledge. [JBH]

Pezzulo, G. (2008) Coordinating with the future: The anticipatory nature of representation. *Minds and Machine* 18:179–25. [GP]

Pezzulo, G. (2013). Why do you fear the Bogeyman? An embodied predictive coding model of perceptual inference. *Cognitive, Affective, and Behavioral Neuroscience.* Available at: http://link.springer.com/article/10.3758%2Fs13415-013-0227-x [GP]

Pezzulo, G. (2011) Grounding procedural and declarative knowledge in sensorimotor anticipation. *Mind & Language* 26:78–114. [GP]

Pezzulo, G. (2012) An active inference view of cognitive control. *Frontiers in Theoretical and Philosophical Psychology* 3:478. doi: 10.3389/fpsyg.2012.00478. [GP]

Pezzulo, G., Baldassarre, G., Butz, M. V., Castelfranchi, C. & Hoffmann, J. (2007) An analysis of the ideomotor principle and TOTE. In: *Proceedings of the Third Workshop on Anticipatory Behavior in Adaptive Learning Systems (ABiALS 2006), Rome, Italy,* ed. M. V. Butz, O. Sigaud, G. Pezzulo & G. Baldassarre, pp. 73–93. Springer-Verlag Berlin, Heidelberg. [GP]

Pezzulo, G. & Castelfranchi, C. (2007) The symbol detachment problem. *Cognitive Processing* 8(2):115–31. [GP]

Pezzulo, G. & Castelfranchi, C. (2009) Thinking as the control of imagination: A conceptual framework for goal-directed systems. *Psychology Research* 73:559–77. [GP]

Pezzulo, G., Rigoli, F. & Chersi, F. (2013) The mixed instrumental controller: Using value of information to combine habitual choice and mental simulation. *Frontiers in Cognition* 4:92. [GP]

Phan, K. L. & Sripada, C. S. (2013) Emotion regulation. In: *The Cambridge handbook of affective neuroscience*, ed.. J. A. A. P. Vuilleumer, pp. 375–400. Cambridge University Press. [CS]

Pichon, S., de Gelder, B. & Grezes, J. (2012) Threat prompts defensive brain responses independently of attentional control. *Cerebral Cortex* 22:274–85. [GBM]

Pinker, S. (1997) *How the mind works.* Norton. [aJYH]

Pinker, S. & Bloom, P. (1990) Natural language and natural selection. *Behavioral and Brain Sciences* 13:707–84. [aJYH]

Popper, K. R. (1972) *Objective knowledge: An evolutionary approach*. Oxford University Press. [aJYH]

Posner, M. J. & Snyder, C. R. (1975) Attention and cognitive control. In: *Information processing in cognition: The Loyola symposium*, ed. R. L. Solso, pp. 55–85. Erlbaum. [aJYH]

Powers, W. T. (1973) *Behavior: The control of perception*. Aldine Publishing Company. [rJYH]

Proffitt, D. R. (2006) Embodied perception and the economy of action. *Perspectives on Psychological Science* 1:110–22. [GP]

Proulx, T., Inzlicht, M. & Harmon-Jones, E. (2012) Understanding all inconsistency compensation as a palliative response to violated expectations. *Trends in Cognitive Sciences* 16:285–91. [ACK]

Pyszczynski, T., Solomon, S., Greenberg, J. & Stewart-Fouts, M. (1995) Liberating and constraining aspects of self: Why the freed bird finds a new cage. In: *The self in European and North American culture: Development and processes*, ed. A. Oosterwegel & R. A. Wicklund, pp. 357–73. Kluwer. [ACK]

Rachlin, H. (1985) Pain and behavior. *Behavioral and Brain Sciences* 8:43–83. [GA]

Raghunathan, R. & Pham, M. T. (1999) All negative moods are not equal: Motivational influences of anxiety and sadness on decision making. *Organizational Behavior and Human Decision Processes* 79:56–77. [EB-M]

Redgrave, P. & Gurney, K. (2006) The short-latency dopamine signal: A role in discovering novel actions? *Nature Reviews Neuroscience* 7(12):967–75. [BE]

Revelle, W. (1995) Personality processes. *Annual Review of Psychology* 46(1):295–328. [EB-M]

Reynolds, K. J., Jones, B., O'Brien, K. & Subasic, E. (in press). System justification and social change in social and political psychology: The importance of subgroup *and* superordinate levels of self-categorization. *European Psychologist*. [ACK]

Robbins, T. W., Ersche, K. D. & Everitt, B. J. (2008) Drug addiction and the memory systems of the brain. *Annals of the New York Academy of Sciences* 1141:1–21. [CPM]

Robinson, T. E. & Kolb, B. (2004) Structural plasticity associated with exposure to drugs of abuse. *Neuropharmacology* 47:33–46. [CPM]

Roe, A. & Simpson, G. G., eds. (1958) *Behavior and evolution*. Yale University Press. [aJYH]

Rogers, C. R. (1951) *Client-centered therapy: Its current practice, implications, and theory*. Hough. [JBH]

Ronay, R. & von Hippel, W. (2010) The presence of an attractive woman elevates testosterone and physical risk-taking in young men. *Social Psychological and Personality Science* 1:57–64. [WvH]

Roskies, A. (2006) Neuroscientific challenges to free will and responsibility. *Trends in Cognitive Science* 10(9):419–23. [NW]

Roskos-Ewoldsen, D. R. & Fazio, R. H. (1992) On the orienting value of attitudes: Attitude accessibility as a determinant of an object's attraction of visual attention. *Journal of Personality and Social Psychology* 63:198–211. [aJYH]

Ross, L. & Nisbett, R. E. (1991) *The person and the situation: Perspectives of social psychology*. McGraw-Hill. [aJYH]

Rupert, R. (2011) Embodiment, consciousness, and the massively representational mind. *Philosophical Topics* 39(1):99–120. [BH]

Russell, J. A. (2003) Core affect and the psychological construction of emotion. *Psychological Review* 110(1):145–72. [EB-M]

Saal, D., Dong, Y., Bonci, A. & Malenka, R. C. (2003) Drugs of abuse and stress trigger a common synaptic adaptation in dopamine neurons. *Neuron* 37:577–82. [CPM]

Saul, J. (2013) Implicit bias, stereotype threat and women in philosophy. In: *Women in philosophy: What needs to change*? ed. F. Jenkins & K. Hutchison, pp. 39–60. Oxford University Press. [NW]

Schacter, D. L. (1992) Implicit knowledge: New perspectives on unconscious processes. *Proceedings of the National Academy of Sciences of the United States of America* 89:11,113–17. [TAM]

Schaller, M. & Neuberg, S. L. (2012) Danger, disease, and the nature of prejudice(s). In: *Advances in experimental social psychology, vol. 46*, ed. J. Olson & M. P. Zanna, pp. 1–55. Academic Press. [SLN]

Schaller, M., Park, J. H. & Mueller, A. (2003) Fear of the dark: Interactive effects of beliefs about danger and ambient darkness on ethnic stereotypes. *Personality and Social Psychology Bulletin* 29:637–49. [SLN]

Scher, S. & Cooper, J. (1989) Motivational basis of dissonance: The singular rule of behavioral consequences. *Journal of Personality and Social Psychology* 56:899–906. [PC]

Schmidt, L., Palminteri, S., Lafargue, G. & Pessiglione, M. (2010) Splitting motivation: Unilateral effects of subliminal incentives. *Psychological Science* 21:977–83. Available at: http://www.ncbi.nlm.nih.gov/pubmed/20511391. [aJYH]

Schmitt, D. P., Alcalay, L., Allik, J., Ault, L., Austers, I., Bennett, K. L., Bianchi, G., Boholst, F., Borg Cunen, M. A., Braeckman, J., Brainerd, E. G. Jr., Caral, L. G. A., Caron, G., Casullo, M. M., Cunningham, M., Daibo, I., De Backer, C., De Souza, E., Diaz-Loving, R., Diniz, G., Durkin, K., Echegaray, M., Eremsoy, E., Euler, H. A., Falzon, R., Fisher, M. L., Foley, D., Fry, D. P., Fry, S., Ghayur, M. A., Golden, D. L., Grammer, K., Grimaldi, L., Halberstadt, J., Herrera, D., Hertel, J., Hoffmann, H., Hooper, D., Hradilekova, Z., Hudek-Kene-evi, J., Jaafar, J., Jankauskaite, M., Kabangu-Stahel, H., Kardum, I., Khoury, B., Kwon, H., Laidra, K., Laireiter, A.-R., Lakerveld, D., Lampert, A., Lauri, M., Lavalleé, M., Lee, S.-J., Chung Leung, L., Locke, K. D., Locke, V., Luksik, I., Magaisa, I., Marcinkeviciene, D., Mata, A., Mata, R., McCarthy, B., Mills, M. E., Moreira, J., Moreira, S., Moya, M., Munyae, M., Noller, P., Opre, A., Panayiotou, A., Petrovic, N., Poels, K., Popper, M., Poulimenou, M., P'yatokha, V., Raymond, M., Reips, U.-D., Reneau, S. E., Rivera-Aragon, S., Rowatt, W. C., Ruch, W., Rus, V. S., Safir, M. P., Salas, S., Sambataro, F., Sandnabba, K. N., Schulmeyer, M. K., Schütz, A., Scrimali, T., Shackelford, T. K., Shaver, P. R., Sichona, F., Simonetti, F., Sineshaw, T., Speelman, T., Spyrou, S., Sümer, H. C., Sümer, N., Supekova, M., Szlendak, T., Taylor, R., Timmermans, B., Tooke, W., Tsaousis, I., Tungaraza, F. S. K., Vandermassen, G., Vanhoomissen, T., Van Overwalle, F., Vanwesenbeeck, I., Vasey, P. L., Verissimo, J., Voracek, M., Wan, W. W. N., Wang, T.-W., Weiss, P., Wijaya, A., Woertman, L., Youn, G. & Zupanèiè, A. (2003) Universal sex differences in the desire for sexual variety: Tests from 52 nations, 6 continents, and 13 islands. *Journal of Personality and Social Psychology* 85:85–104. [DC-B]

Schröder, T. & Thagard, P. (2013) The affective meanings of automatic social behaviors: Three mechanisms that explain priming. *Psychological Review* 120 (1):255–80. [BH]

Schwarz, N. & Bohner, G. (1996) Feelings and their motivational implications: Moods and the action sequence. In: *The psychology of action: Linking cognition and motivation to behavior*, ed. P. M. Gollwitzer & J. A. Bargh, pp. 119–45. Guilford. [EB-M]

Seymour, B. & McClure, S. M. (2008) Anchors, scales and the relative coding of value in the brain. *Currrent Opinion in Neurobiology* 18(2):173–78. Available at: http://www.ncbi.nlm.nih.gov/pubmed/18692572. [CS]

Shah, J. Y. (2003) Automatic for the people: How representations of significant others implicitly affects goal pursuit. *Journal of Personality and Social Psychology* 84:661–81. Available at: http://psycnet.apa.org/journals/psp/84/4/661/. [aJYH, GBM]

Shah, J. Y., Friedman, R. & Kruglanski, A. W. (2002) Forgetting all else: On the antecedents and consequences of goal shielding. *Journal of Personality and Social Psychology* 83:1261–80. Available at: http://www.ncbi.nlm.nih.gov/pubmed/12500810. [arJYH, AF, CK]

Shah, J. Y. & Gardner, W. L. (2008) *Handbook of motivation science*. Guilford. [GBM]

Shapiro, J. A. (2009) Revisiting the central dogma in the 21st century. *Annals of the New York Academy of Sciences* 1178:6–28. doi: 10.1111/j.1749-6632.2009.04990.x. [TAM]

Shapiro, L. (2007) The embodied cognition research programme. *Philosophy Compass* 2(2):338–46. [NW]

Sharot, T., Korn, C. W. & Dolan, R. J. (2011) How unrealistic optimism is maintained in the face of reality. *Nature Neuroscience* 14:1475–79. [GP]

Sheeran, P., Gollwitzer, P. M. & Bargh, J. A. (2013) Nonconscious processes and health. *Health Psychology* 32:460–73. [aJYH]

Sheldon, K. M. & Kasser, T. (1995) Coherence and congruence: Two aspects of personality integration. *Journal of Personality and Social Psychology* 68(3):531–43. [JBH]

Sherif, M., Harvey, O. J., White, B. J., Hood, W. R. & Sherif, C. (1961) *Intergroup conflict and cooperation: The robbers cave experiment*. The University Book Exchange. [arJYH]

Shiffrin, R. M. & Schneider, W. (1977) Controlled and automatic human information processing: II. Perceptual learning, automatic attending, and a general theory. *Psychological Review* 84:127–90. [aJYH]

Shizgal, P. (2012) Scarce means with alternative uses: Robbins' definition of economics and its extension to the behavioral and neurobiological study of animal decision making. *Frontiers in Neuroscience* 6(20). doi: 10.3389/fnins.2012.00020. [DS]

Shultz, W., Dayan, P. & Monague, R. (1997) A neural substrate of prediction and reward. *Science* 275:1593–99. [BH]

Silvestrini, N. & Gendolla, G. H. (2013) Automatic effort mobilization and the principle of resource conservation: One can only prime the possible and justified. *Journal of Personality and Social Psychology* 104:803–16. [BE]

Simon, H. A. (1956) Rational choice and the structure of the environment. *Psychological Review* 63:129–38. [aJYH]

Simon, L., Greenberg, J. & Brehm, J. (1995) Trivialization: The forgotten mode of dissonance reduction. *Journal of Personality and Social Psychology* 68:247–60. [PC]

Simons, D. J. & Chabris, C. F. (1999) Gorillas in our midst: Sustained inattentional blindness for dynamic events. *Perception* 28:1059–74. Available at: http://wexler.free.fr/library/files/simons%20(1999)%20gorillas%20in%20our%20midst.%20sustained%20inattentional%20blindness%20for%20dynamic%20events.pdf. [aJYH]

Sinclair, L. & Kunda, Z. (2000) Motivated stereotyping of women: She's fine if she praised me but incompetent if she criticized me. *Personality and Social Psychological Bulletin* 26:1329–42. Available at: http://psp.sagepub.com/cgi/reprint/26/11/1329. [aJYH]

Skinner, B. F. (1953) *Science and human behavior*. Macmillan. [aJYH]

Smith, E. E. & Jonides, J. (1998) Storage and executive processes in the frontal lobes. *Science* 283:1657–61. Available at: http://www.sciencemag.org/cgi/content/full/283/5408/1657?ijkey=HiJyaWahcPzIA. [aJYH]

Solway, A. & Botvinick, M. M. (2012) Goal-directed decision making as probabilistic inference: A computational framework and potential neural correlates. *Psychology Review* 119:120–54. [GP]

Spencer, S. J., Fein, S., Wolfe, C., Fong, C. & Dunn, M. (1998) Automatic activation of stereotypes: The role of self-image threat. *Personality and Social Psychology Bulletin* 24:1139–52. Available at: http://www.asc.upenn.edu/usr/ogandy/C53704read/automaticactivation.pdf. [aJYH]

Sripada, C. S. (2014) How is willpower possible? The puzzle of synchronic self-control and the divided mind. *Noûs* 48:41–74. [CS]

Srull, T. K. & Wyer, R. S., Jr. (1979) The role of category accessibility in the interpretation of information about persons: Some determinants and implications. *Journal of Personality and Social Psychology* 37:1660–72. [aJYH]

Stearns, S. C. (2012) Evolutionary medicine: Its scope, interest and potential. *Proceeding of the Royal Society, London, Series B* 279:4305–21. [BM]

Stopford, J. M. (2003) Organizational learning as guided responses to market signals. In: *Handbook of organizational learning and knowledge*, ed. M. Dierkes, A. B. Antal, J. Child & I. Nonaka, pp. 264–81. Oxford University. [rJYH]

Strack, F. & Deutsch, R. (2004) Reflective and impulsive determinants of behavior. *Personality and Social Psychology Review* 8:220–47. Available at: https://www.uni-wuerzburg.de/fileadmin/uniwue/_temp_/02strack_und_deutsch.pdf. [aJYH]

Strack, F. & Deutsch, R. (in press) The duality of everyday life: Dual-process and dual-system models in social psychology. In: *Handbook of personality and social psychology, vol. 1, attitudes and social cognition,* ed. E. Borgida & J. Bargh. American Psychological Association. [aJYH]

Stroebe, W., Mensink, W., Aarts, H., Schut, H. & Kruglanski, A. W. (2008) Why dieters fail: Testing the goal conflict model of eating. *Journal of Experimental Social Psychology* 44:26–36. Available at: http://www.goallab.nl/publications/documents/Stroebe,%20Mensink,%20Aarts,%20Schut,%20Kruglanski%20(2008)%20-%20why%20dieters%20fail.pdf. [aJYH]

Suhler, C. L. and Churchland, P. S. (2009) Control: conscious and otherwise. *Trends in Cognitive Sciences* 13:341–47. [aJYH]

Sullivan, D., Landau, M. J. & Kay, A. C. (2012) Toward a comprehensive understanding of existential threat: Insights from Paul Tillich. *Social Cognition* 30:734–57. [ACK]

Swain, J. E. (2011) The human parental brain: In vivo neuroimaging. *Progress in Neuropsychopharmacology and Biological Psychiatry* 35(5):1242–54. Available at: http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=21036196. [CS]

Swain, J. E., Kim, P., & Ho, S. S. (2011) Neuroendocrinology of parental response to baby-cry. *Journal of Neuroendocrinology* 23(11):1036–41. [CS]

Swain, J. E., Mayes, L. C., & Leckman, J. F. (2004) The development of parent-infant attachment through dynamic and interactive signaling loops of care and cry. *Behavioral and Brain Sciences* 27(4):472–73. [CS]

Symons, D. (1979) *The evolution of human sexuality*. Oxford University Press. [DC-B]

Symons, D. (1992) On the use and misuse of Darwinism in the study of human behavior. In: *The adapted mind: Evolutionary psychology and the generation of culture*, ed. J. H. Barkow, L. Cosmides & J. Tooby, pp. 137–59. Oxford University Press. [aJYH, DC-B]

Tetlock, P. E. (2002) Social functionalist frameworks for judgments and choice: Intuitive politicians, theologians, and prosecutors. *Psychological Review* 109:451–71. Available at: http://www.ncbi.nlm.nih.gov/pubmed/12088240. [aJYH]

Thaler, R. H. & Shefrin, H. M. (1981) An economic theory of self-control. *Journal of Political Economy* 89:392–406. [AF]

Tomasello, M., Carpenter, M., Call, J., Behne, T. & Moll, H. (2005) Understanding and sharing intentions: The origins of cultural cognition. *Behavioral and Brain Sciences* 28:675–91. Available at: http://email.eva.mpg.de/~tomas/pdf/BBS_Final.pdf. [aJYH]

Tooby, J. & Cosmides, L. (1992) The psychological foundations of culture. In: *The adapted mind: Evolutionary psychology and the generation of culture*, ed. J. H. Barkow, L. Cosmides & J. Tooby, pp. 19–136. Oxford University Press. [aJYH, DC-B]

Ungless, M. A., Whistler, J. L., Malenka, R. C. & Bonci, A. (2001) Single cocaine exposure in vivo induces long-term potentiation in dopamine neurons. *Nature* 411:583–87. [CPM]

Van Bavel, J. J. & Cunningham, W. A. (2009) Self-categorization with a novel mixed-race group moderates automatic social and racial biases. *Personality and Social Psychology Bulletin* 35:321–35. doi: 10.1177/0146167208327743. [aJYH]

Van Koningsbruggen, G. M., Stroebe, W. & Aarts, H. (2011) Through the eyes of dieters: Biased size perception of food following tempting food primes. *Journal of Experimental Social Psychology* 47:293–99. [GBM]

Vaughan, W., Jr. & Herrnstein, R. J. (1987) Stability, melioration, and natural selection. In: *Advances in behavioral economics, vol. 1,* eds. L. Green & J. H. Kagel, pp. 185–215. Ablex. [GA]

Veling, H. & van Knippenberg, A. (2006) Shielding intentions from distraction: Forming an intention induces inhibition of distracting stimuli. *Social Cognition* 24:409–25. Available at: http://www.goallab.nl/publications/documents/Veling,%20van%20Knippenberg%20%282006%29%20-%20shielding%20intentions%20from%20distraction.pdf. [aJYH]

Veltkamp, M., Aarts, H. & Custers, R. (2008) Perception in the service of goal pursuit: Motivation to attain goals enhances the perceived size of goal-instrumental objects. *Social Cognition* 26:720–36. Available at: http://goallab.nl/publications/documents/Veltkamp,%20Aarts,%20Custers%20%28in%20press%29%20Social%20Cognition.pdf. [arJYH]

Verschoor, S. A., Spapé, M., Biro, S. & Hommel, B. (2013) From outcome prediction to action selection: Developmental change in the role of action-effect bindings. *Developmental Science* 16:801–14. [BE]

Verschure, P. F. M. J. (2012) Distributed adaptive control: A theory of the mind, brain, body nexus. *Biologically Inspired Cognitive Architectures* 1:55–72. [GP]

Vogeley, K. & Fink, G. R. (2003) Neural correlates of the first-person-perspective. *Trends in Cognitive Sciences* 7:38–42. [TAM]

Vohs, K. D., Mead, N. L. & Goode, M. R. (2006) The psychological consequences of money. *Science* 314:1154–56. Available at: http://carlsonschool.umn.edu/Assets/71704.pdf. [aJYH]

Volkow, N. D. & Fowler, J. S. (2000) Addiction, a disease of compulsion and drive: Involvement of the orbitofrontal cortex. *Cerebral Cortex* 10:318–25. [NLC]

von Hippel, W. & Trivers, R. (2011) The evolution and psychology of self-deception. *Behavioral and Brain Sciences* 34:1–16. [aJYH]

von Schelling, F. (1800/1978) *System des transcendentalen Idealismus/System of Transcendental Idealism,* trans. P. Heath. University Press of Virginia. [TAM]

Webb, T. L. & Sheeran, P. (2006) Does changing behavioral intentions engender behavior change? A meta-analysis of the experimental evidence. *Psychological Bulletin* 132:249–68. [aJYH]

Wegener, D. T., Petty, R. E. & Smith, S. M. (1995) Positive mood can increase or decrease message scrutiny: The hedonic contingency view of mood and message processing. *Journal of Personality and Social Psychology* 69:5–15. [EB-M]

Wegner, D. M. (2002) *The illusion of conscious will*. MIT Press. [aJYH]

West-Eberhard, M. J. (2003) *Evolution and developmental plasticity*. University Press. [MF]

Wexler, M., Kosslyn, S. & Berthoz, A. (1998) Motor processes in mental rotation. *Cognition* 68:77–94. [BH]

White, N. M. (1996) Addictive drugs as reinforcers: Multiple partial actions on memory systems. *Addiction* 91:921–49. [CPM]

Whitman, W. (1855/1960) *Leaves of grass*. Viking Press. (Original work published in 1855.) [aJYH]

Wicklund, R. A. & Gollwitzer, P. M. (1982) *Symbolic self-completion.* Erlbaum. [GBM]

Wiers, R. W., Houben, K., Roefs, A., de Jong, P., Hofmann, W. & Stacy, A. W. (2010) Implicit cognition in health psychology: Why common sense goes out the window. In: *Handbook of implicit social cognition*, ed. B. Gawronski & B. K. Payne, pp. 463–88. Guilford. [aJYH]

Williams, L. A. & DeSteno, D. (2014) Building bonds, attaining ambitions, and establishing esteem: How positive emotions serve intrapersonal needs. In: *The light and dark sides of positive emotion*, ed. J. Gruber & J. Moskowitz, pp. 206–24. Oxford University Press. [EB-M]

Wilson, T. D. (2002) *Strangers to ourselves: Discovering the adaptive unconscious*. Harvard University Press. [TAM]

Wilson, T. D. & Brekke, N. (1994) Mental contamination and mental correction: Unwanted influences on judgments and evaluations. *Psychological Bulletin* 116:117. [aJYH]

Yik, M. S. M. & Russell, J. A. (2001) Predicting the big two of affect from the Big Five of personality. *Journal of Research in Personality* 35(3):247–77. [EB-M]

Zeigarnik, B. (1927) Das Behalten erledigter und unerledigter Handlungen [The retention of completed and uncompleted actions]. *Psychologische Forschung* 9:1–85. [GBM]