Unconscious Behavioral Guidance Systems

John A. Bargh
Department of Psychology, Yale University

Ezequiel Morsella
Department of Psychology, San Francisco State University;
Department of Neurology, University of California, San Francisco


Authors’ Note. Preparation of this chapter was supported in part by Grant R01-MH60767 from the U.S. National Institute of Health. Thanks to Don Carlston for helpful feedback on an earlier version.
It is the duty of the natural scientist to attempt a natural explanation before he contents himself with drawing upon factors extraneous to nature.

—Konrad Lorenz (1962, p. 23)

Introduction

In the early days (by which we mean way way way back in the 1980s), discoveries of nonconscious processes were seen as magical, mysterious -- if they were believed at all. The conscious-process account of how one gets from A to B (say, from attitudes to behavior, or from a witnessed behavior to one’s attribution of its cause) involved self-reportable steps or stages, clear and easy for all to see. But when the same effects began to be demonstrated without conscious awareness or involvement, the underlying process was invisible, and quite mysterious (especially mysterious, for some reason, to journal editors and reviewers!). Happily, in the three decades that followed, our understanding of how these nonconscious or automatic processes operate has improved a hundredfold (see Bargh, 2006). No longer are these effects viewed as miraculous; today they are just as theoretically tractable as conscious or controlled processes, and appeals to divine miracles are no longer necessary to explain them.

Purposive behavior is widespread among living organisms, although only a few possess anything approaching the information processing capacities of human consciousness (Mayr, 1976; Tomasello et al., 2005). The behavior of most organisms living today—the fly, the venus fly trap, and perhaps the alligator and its fellow reptiles—is under the guidance of what we humans consider to be unconscious control. (For a treatment regarding the presence of consciousness in non-human animals, see Gray, 2004.) Moreover, for millions of years prior to the advent of consciousness, the actions of intelligent life forms, both toward the world and toward each other, was under the control of exclusively unconscious systems. Conscious mental processes (e.g., conscious intentions) were
relatively late comers in the phylogeny of intelligent behavior (Corballis, 2007; Deacon, 1997; Dennett, 1991; Donald, 1991). Thus, as difficult as it may be for us to imagine—because consciousness encompasses the totality of what ‘we’ are and could ever experience—the history of our planet and its organisms was for the most part one devoid of consciousness: A zombie-like world having colorful plants, streams, and gigantic creatures performing complex acts and communicating with each other, but possessing nothing that could ever be conscious of any of it.

Hence, despite our intuitions and how much consciousness may mean to ‘us,’ consciousness is actually an atypical phenomenon and tool with respect to both the natural world and the majority of human nervous function. Figuratively speaking, it is as atypical regarding the nuts and bolts of intelligent behavior as are the computerized, GPS-based navigational systems in today’s automobiles. These devices, and their ‘reverse engineering,’ fail to reveal the basic principles of mechanized transport, and, as sophisticated as they may be (interacting with satellites and creating graphical ‘representations’ of one’s current driving environment, including traffic conditions), they are not responsible for either powering the car nor for conducting it. Just as automobiles could get from one place to another long before the advent of such systems, so did creatures express intelligent behavior before the advent of consciousness. The intelligentia of this ‘unconscious fauna’ is still within us, and, like the engine and driver of a car, it is working behind the scenes as the prime mover of our behavioral repertoire, in the form of unconscious behavioral guidance systems.

Who is the real driver behind the wheel of the cognitive apparatus?

Because evolution works only gradually, making incremental changes and, if at all possible, “exapting” already-existing structures and processes (Allman, 2000; Bargh & Morsella, 2008), we should be able to find evidence still today of the continued operation of these original unconscious behavioral guidance systems. How would such systems work? The simplest mechanism, present in
all living organisms from single-cell paramecia to human beings, is reflexive approach versus withdrawal behaviors, in direct response to external stimuli (Schneirla, 1959). For many animals, incoming stimuli would activate approach or withdrawal behaviors (Roe & Simpson, 1958). These are simple reflex or S-R (stimulus-response) reactions. Humans of course are much more complex information processors than this, with sophisticated internal processing systems specializing in producing affective, emotional, cognitive, perceptual, and motivational responses to external stimuli, each of which mediate between the sensation of a stimulus event and its effect on behavior (Bargh, 1997). Despite the claims of the radical behaviorists (e.g., Skinner, 1953; see Bargh & Ferguson, 2000), human behavior is not controlled by such simple S-R, direct stimulus-to-behavior linkages: indeed, the concept of the simple reflex arc on which behaviorist theory was built had long ago been abandoned by physiologists such as Sherrington (1906), who called it “a convenient, if not probable, fiction” (p. 137). Skinner’s (1957) attempt to account for the higher mental processes (including language and social interaction) in humans using only such simple S-R associations failed spectacularly, and helped in fact to bring about the cognitive revolution (Chomsky, 1959; Koestler, 1967), as he permitted himself no recourse to internal systems to extract the important, situationally- and purpose-relevant meanings from external stimuli, and to prepare the individual for appropriate behavioral responses.

Research on priming and automaticity in social psychology, on the other hand, has examined the direct (unconscious) effect of environmental stimuli on these important internal mediational systems, and has shown that the mere, passive perception of environmental events directly triggers higher mental processes in the absence of any involvement by conscious, intentional processes (see reviews in Bargh & Ferguson, 2000; Dijksterhuis, Aarts, & Chartrand, 2007; Higgins, 1996). These automatic effects of environmental stimuli were found to drive evaluation (e.g., Fazio, 1986, 1990),
stereotyping and prejudice (Devine, 1989), social behavior (e.g., Bargh, Chen, & Burrows, 1996; Dijksterhuis & van Knippenberg, 1998), and motivated goal pursuit (e.g., Bargh & Gollwitzer, 1994; Chartrand & Bargh, 1996) – in each case, without any awareness by the individual of the role played by these external stimuli in the production of his or her behavior. The ease and ubiquity with which priming effects on these sophisticated higher mental (including executive) processes have been obtained reveals both the openness of the human mind to environmental influences, and a necessarily decreased role for intentional, conscious causation and guidance of the higher mental processes (Bargh & Ferguson, 2000; Huang & Bargh, 2008).

*The present model*

According to the above logic, the human version of unconscious behavioral guidance systems must have (at minimum) two main stages, not just one (i.e., S-R): the initial automatic activation of the mediating system by external stimuli (*Step 1*) and that system’s effect on behavior (*Step 2*). Both stages must be capable of unconscious operation; that is, with no role played by conscious choice or guidance in the entire sequence, for us to be able to speak of truly unconscious behavioral guidance systems (see Figure 1).
What are the different mental systems that are directly and unconsciously activated by external stimuli? Research has identified four: perceptual, evaluative, motivational, and emotional. They are considered distinct systems because they are dissociable – they have different operating characteristics and qualities and are not reducible to each other. For example, activations in the

---

1 How researchers define the unconscious significantly affects conclusions as to its power over human behavior and other higher mental processes (Bargh & Morsella, 2008). In cognitive science, the powers of the unconscious are often operationally equated with the powers of subliminally-presented stimuli (i.e., one must be unaware of the triggering stimulus itself). This practice has led to the conclusion that the unconscious is rather “dumb” (Loftus & Klinger, 1992), because while concept activation and primitive associative learning could occur unconsciously, nothing complex requiring flexible responding, integration of stimuli, or higher mental processes could. But this was not the original meaning of the term “unconscious”, which was used by Darwin (1859), Freud (see Brill, 1938), and others to refer to the unintentional nature of the behavior or process, with an associated lack of awareness not of the stimuli themselves, but of the influence or consequences of those stimuli. It is this “unintentional” definition that has driven research on unconscious (automatic, nonconscious, implicit) phenomena in social psychology since Nisbett and Wilson (1977) posed their seminal question: “To what extent are people aware of and able to report on the true causes of their behavior?”
perceptual system are relatively short lived, and those in the evaluative system are even shorter, but activated motivational representations (goal pursuit structures) actually *increase* in strength over time (until the goal is actively pursued: Atkinson & Birch, 1971; Bargh, Gollwitzer, Lee-Chai, Barndollar, & Troetschel, 2001, Study 2).

In what follows we will first sketch out the evidentiary support for these four unconscious behavior guidance systems; within each of the four types we present the basic evidence for Step 1 priming or activation effects, followed by the Step 2 evidence that these priming effects extend to the creation of appropriate behavioral tendencies.

**Unconscious Behavioral Guidance Systems**

*Perceptual*

*Theoretical background.* Recent cognitive neuroscience research supports the dissociation of conscious awareness and intention from the operation of complex behavioral processes (see Prinz, 2003). First, in a classic study by Goodale, Milner, Jakobsen, and Carey (1991), patients with lesions in the parietal lobe region are able to correctly identify an object held up to them by an experimenter, but are not able to reach for it correctly based on its spatial orientation (horizontal or vertical). Patients with lesions in the ventral-visual system, on the other hand, could not identify (recognize) the object, but were still able to reach for it correctly when the experimenter casually asks them to take it. Thus, the one group exhibited appropriate action tendencies toward the object in the absence of conscious awareness of what the object was (i.e., action without perception), while the other group was aware of what the object was but could not act towards it appropriately (i.e., perception without action). Theorists have concluded from this and related studies that two different cortical visual pathways are activated in the course of perception, a dorsal pathway that supports actional responses (“what to do”) and a ventral pathway supporting semantic knowledge regarding
the object (“what it is”; see review in Westwood, 2009). Importantly, it is the dorsal (actional) system that is believed to operate largely outside of conscious awareness, while the operation of the ventral system is normally accessible to awareness (Decety & Grèzes, 1999; Jeannerod, 2003; Norman, 2002).

Indeed, humans are generally unaware of the operations of their motor system (Fourneret & Jeannerod, 1998; Frith et al., 2000; Rosenbaum, 2002). This research is revealing a startling lack of awareness on the part of individual human beings of exactly how one is moving one’s body through space. Fourneret and Jeannerod (1998) showed that when one’s hand is controlling a computer-drawing device but behind a screen so the individual is prevented from seeing the hand in motion, participants can be easily fooled into thinking their hand moved one direction when it had actually moved in a different direction (through false feedback on the computer display). Participants reported great confidence that their hand had moved in the direction of the line drawn on the screen, when in reality substantial bias had been programmed into the translation of their actual movement into what was displayed. This result is obtainable only if participants had little if any conscious access to their actual hand movements.

One important function of social perception of which people are generally unaware is its direct effect in preparing one’s own behavioral responses. The priming effects of people’s behavior and other situational features extend to a direct influence on our own behavior, beginning soon after birth. Infants naturally learn much about how to behave by mere passive imitation of fellow children and also their adult caretakers. Meltzoff (2002) concluded from decades of researching this phenomenon that infants can imitate body movements and facial acts at birth, and that this ability represents a “primordial connection between infant and caretaker” (p. 19).
That infants engage in direct imitation of others’ behavior indicates that such imitation is likely not a strategic or intentional act on their part, but the outcome of an unconscious tendency to act in harmony with those around us. As Dawkins (1976) pointed out, the best behavioral strategy from the point of view of evolution and adaptation “depends on what the majority of the population is doing” (p. 69; see Maynard Smith, 1982; Maynard Smith & Parker, 1976). Thus “blindly” or unconsciously adopting what others around you are doing, especially in new situations or with strangers, makes good adaptive sense as a default option or starting point for your own behavior.

Step 1. Automatic or unconscious influences on social perception were among the first experimental demonstrations of automaticity in social information processing. Individuals as social objects are automatically categorized in terms of their social groups (race, gender, roles, etc.), and social behaviors are automatically categorized in trait terms (e.g., Uleman et al., 1996), so that social perceivers naturally “go beyond the information given” (e.g., Brewer, 1988; Bruner, 1957; Fiske, 1982). The preconscious analysis of the environment that is continually occurring during normal perceptual activity (e.g., Neisser, 1967) was found, in the case of social information, to extend to a rich activation of knowledge, assumptions and expectancies concerning the individuals one encounters, much of it at an implicit or unconscious level (as in automatic stereotype activation; e.g., Devine, 1989). The ease and ubiquity of priming effects in social psychology (and now in other fields such as behavioral economics, health psychology, consumer behavior, and political psychology) is testimony to the rich variety of meanings unconsciously activated during normal perceptual activity (see reviews in Bargh, 1989, 2007; Higgins, 1996).

Step 2. These imitative impulses, triggered by the perceived behavior of others, continue to be activated throughout one’s life, causing children and adults to have default tendencies to act the same as those around them are acting – producing behavioral and emotional contagion effects.
Thus, how other people are acting around us in the present is yet a further unconscious influence or guide as to how we ourselves should act. This tendency, and its unconscious and unintentional nature, has been repeatedly demonstrated in human adults in the research of Chartrand and colleagues (e.g., Chartrand & Bargh, 1999; Chartrand, Maddux, & Lakin, 2005; Lakin & Chartrand, 2003). People don’t know and even don’t believe once informed that they had engaged in these imitative behaviors. Not only do people tend to adopt the physical behavior (posture, facial gestures, arm and hand movements) of strangers with whom they interact, without intending to or being aware they are doing so, such unconscious imitation also tends to increase liking and bonding between the individuals – serving as a kind of natural ‘social glue’ (see also Giles, Coupland, & Coupland, 1991).

Wiltermuth and Heath (2009) recently extended this finding from the dyadic to the group level. They pointed to the widespread use of synchronous rituals by social groups over thousands of years of human history, and argued that these rituals serve the purpose of promoting cohesiveness within the group: for example, armies marching in lockstep, and worshipers at religious services standing, kneeling, and singing in unison. They provided experimental demonstrations of this natural phenomenon by first manipulating synchronous activity – for example, by having participants walk in step with each other (or not) on a stroll around campus, mouthing song lyrics together (or not) while listening to a well-known song. Compared to those in the control group, those who acted in synchrony with each other then were found to cooperate with each other to a greater extent across a variety of cooperative tasks. Doing the same things at the same time as one’s group-mates fosters bonding and cooperative behavior within the group, just as it does within dyads.

Finally, just as the internal meanings that are automatically activated during social perception (Step 1) extend beyond what is directly observable in the current environment (as in stereotype or trait concept activations), so too do the automatic behavioral effects (Step 2) of those perceptual
activations (see reviews in Dijksterhuis & Bargh, 2001; Dijksterhuis, Aarts, & Chartrand, 2007). Stereotypic content activated during social perception creates behavioral tendencies to act in line with that content, across a wide range of stereotypes and types of behavior studied (e.g., activating the elderly stereotype influences both physical behavior such as walking speed but also cognitive performance such as forgetfulness). The perception of situations also automatically activates behavioral dispositions in line with the particular setting; for example, priming the concept of ‘library’ causes participants (not actually in a library) to speak more quietly (Aarts & Dijksterhuis, 2002). Because the individual is unaware of both the stereotype activation and its influence on action tendencies, these studies show a direct, unconscious influence of the current environment on behavior that is mediated by mundane perceptual activity.

Related evidence. The discovery of mirror neurons in both macaques (Rizzolatti & Arbib, 1998) and humans (Buccino et al., 2001), in which perception of behavior in a conspecific directly (unconsciously) activates the same premotor cortical regions used to engage in that behavior oneself, is significant further support for the proposed direct, unconscious influence of perception on behavior. This tight, automatic connection between our perceptual and our actional representations suggests that we are prewired to have behavioral and goal-pursuit tendencies in line with those around us.

Moreover, related cognitive neuroscience research has indicated as well an automatic connection between behavioral concept representations and their corresponding motor representations. Merely hearing action verbs pronounced out loud activates the same brain region (Brodman 45) as does witnessing a meaningful action (Jeannerod, 1999) and both activate implicit motor representations needed to carry out that type of behavior (Perani et al., 1999). Motor programs thus appear to be part of the very meaning of action-related verbs (Grèzes & Decety, 2001;
Pulvermuller, 2005), and this fact is likely responsible for the many successful behavioral priming demonstrations using verbal stimuli (as in the commonly used Scrambled Sentence Test measure of priming).

Evaluative

Theoretical background. Preferences and feelings are unconscious guides to appropriate behavior. A tight connection between immediate, unconscious evaluation and appropriate (approach versus avoidance) actional tendencies is found throughout the animal kingdom; even single-celled paramecia have them (Schneirla, 1959). These “guides” do not arise out of thin air, however. Our present preferences are derived from those that served adaptive ends in the past. Knowledge gained at a lower level of blind selection, the short-cuts and other “good tricks” (Dennett, 1995) that consistently worked over our long-term evolutionary past, are fed upwards as a starting point—appearing as a priori knowledge, the source of which we are unaware. Campbell (1974) called these “shortcut processes” because they save us from having to figure out, each of us individually from scratch, what are the good and helpful things and which are the dangerous.

Step 1. The finding that attitudes can become active automatically upon the perception of the attitude object (Fazio, Sanbonmatsu, Powell, & Kardes, 1986; see also Fiske, 1982) was one of the first demonstrations of automaticity in social psychology. The mere presentation of an attitude object name (prime) was shown to automatically and immediately activate its associated evaluation (good versus bad), in that the activated evaluation facilitated or interfered with evaluation of a second presented attitude object (target) in a Stroop-like fashion. Subsequent research using this and other experimental techniques (e.g., the Implicit Association Test; Nosek et al., 2007) revealed automatic evaluation to be a fairly general and ubiquitous phenomenon, occurring even for novel
stimuli (e.g., fragments of abstract art) the participant had not encountered before (Duckworth, Bargh, Garcia, & Chaiken, 2001).

**Step 2.** Under the present argument that the unconscious evolved as a behavioral guidance system, a source of adaptive and appropriate actional impulses, these unconsciously activated preferences should be found to be directly connected to behavioral mechanisms. Several studies have now established this connection: immediate and unintended evaluation processes are directly linked to approach and avoidance behavioral predispositions in humans, as theorized originally by both Lewin (1935) and Osgood (1957). Chen and Bargh (1999; see also Kawakami et al., 2000; Neumann, Förster, & Strack, 2003) showed that participants are faster to make approach movements of the arm (pulling a lever towards oneself) when responding to positive attitude objects, and faster to make avoidance movements (pushing the lever away) when responding to negative attitude objects. This was true even though their conscious task in the experiment was not to evaluate the objects at all, merely to “knock off the screen” the names of these objects as soon as they appeared. And this unconscious behavioral tendency to approach what is good and avoid what is bad extends even to novel objects never encountered before; Duckworth et al. (2002) obtained the evaluation-action link for the novel as well as the mundane attitude objects in their study.

**Related evidence.** This “evaluation-motivation-action” effect appears to be bi-directional: Cacioppo, Priester, and Berntson (1993) found that participants induced into approach-related versus avoidance-related muscular movements while being exposed to novel attitude objects caused evaluations of the objects (i.e., attitude formation) to be consistent with the induced motivational orientation. Recently, van Knippenberg et al. (2007) showed that inducing participants to make avoidance-related (versus approach-related) muscular movements resulted in more effortful cognitive processing on a subsequent task, in line with now-established findings that negative stimuli
attract greater attention (Smith, Cacioppo, Larsen, & Chartrand, 2003) than positive stimuli, and negative moods recruit effortful processing to a greater extent than positive moods (e.g., Schwarz & Clore, 1996).

Motivational

Theoretical background. Goals can be conceptualized as mental representations of desired end-states that include the means through which to attain those states (Aarts & Dijksterhuis, 2000; Bargh, 1990; Hassin, in press; Kruglanski et al., 2002; Fishbach & Ferguson, 2007; McCulloch, Ferguson, Kawada, & Bargh, 2008). Theoretically, then, goals can thus be primed, or unconsciously activated by relevant environmental stimuli, just as can other representations (Bargh, 1990).

That a goal can operate independently of conscious awareness of its operation implies the existence of a dissociation between the executive control structures in the brain responsible for ‘running’ that goal’s ‘program’ and those that enable conscious awareness of the goal pursuit. Recent cognitive neuroscience research has confirmed that distinct anatomical structures support the operating goal program, on the one hand, and the knowledge of its operation (i.e., consciously-held intentions) on the other. As one review concluded, aspects of the processing of conscious intentions appear to be represented in the prefrontal and premotor cortex, but it is the parietal cortex that houses the representation used to guide action (Frith, Blakemore, & Wolpert, 2000).

Moreover, the hypothesis that executive control structures could operate without the person’s awareness of their operation requires 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 had been reported in stroke patients with ‘environmental dependency syndrome’ (Lhermitte, 1986; also Bogen, 1995), whose behavior was almost entirely driven by situational cues. This hypothesis is supported as well by
research in which incentives associated with various goal pursuits were manipulated outside of participants’ awareness and nonetheless influenced effort expenditure at the task (Aarts, Custers, & Marien, 2008; Pessiglione et al., 2007).

Step 1. Research on unconscious goal pursuit has typically primed a goal representation and then measured its predicted effect on behavior. However, this is the same method used to demonstrate perception-behavior effects, and so from the behavioral effects alone it can be difficult to tell if the responsible mechanism is an activated motivational state or a straight, nonmotivational activation of a behavioral predisposition (see discussion in Dijksterhuis & Bargh, 2001). Therefore, other means must be found in order to be able to tell one system’s operation from the other.

Fortunately, the pioneering theoretical and empirical work of Kurt Lewin (1926, 1935) provided us with signature qualities of motivational states that permit us to classify an effect as motivational instead of cognitive, evaluative, or perceptual. Classic research on conscious goal pursuit confirmed the existence of the signature phenomenal qualities of motivational as posited by Lewin (Bandura, 1977, 1986; Gollwitzer & Moskowitz, 1996; Heckhausen, 1991; Lewin, 1926), and more recent research has confirmed that these qualities hold as well for unconscious goal pursuit. These qualities include persistence in the face of obstacles, resumption of interrupted goal pursuits despite the availability of intrinsically more attractive activities (Bargh et al., 2001; Chartrand & Bargh, 2002; Custers, Maas, Wildenbeest, & Aarts, 2008; Ferguson, Hassin, & Bargh, 2008), increase in goal strength over time (Atkinson & Birch, 1971; Bargh et al., 2001, Study 2), and changes in mood and goal strength depending on the success versus failure of the goal attempt (Bongers, Aarts, & Dijksterhuis, 2009; Chartrand & Bargh, 2002).

Cognitive neuroscience studies of the brain regions involved in motivated behavior support the hypothesis that the same underlying mechanisms are implicated in unconscious as in conscious
goal pursuit. Pessiglione et al. (2007) showed an automatic effort increase effect in response to increases in incentive or reward on a hand-grip exertion task, both when the reward cue (amount of money to be won on that trial) was presented to conscious awareness as when it was presented subliminally. They also found that the same region of the basal forebrain moderates task effort level in response to both the consciously perceived and the subliminally presented reward signal. 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” (p. 906). Neurophysiological recordings also show that the same brain regions are invoked whether the goal pursuit is conscious or unconscious (Pessiglione et al., 2007).

Moreover, just as perceptual-priming effects on social behavior are relatively short-lived and decrease over time (e.g., Higgins, Bargh, & Lombardi, 1985), perceptually-driven unconscious effects on behavior decrease over time as well, but goal or motivation priming effects actually increase over time (Atkinson & Birch, 1971) until the goal is pursued. Bargh et al. (2001; Study 2) demonstrated this dissociation of these two effects over a time delay: effects of an achievement prime on social perception (impression formation) decreased and disappeared after a 5 min delay, but its effects on task performance behavior increased over the same time interval. Taken together, research has revealed unconscious goal pursuit to produce the same outcomes, through the same underlying processes and involvement of the same brain regions, and to possess the same phenomenal qualities as in conscious goal pursuit (see Huang & Bargh, 2008), supporting the present conclusion that goal representations were indeed automatically activated (Step 1) and responsible for the behavioral effects described next under “Step 2”.

**Step 2.** The goal-priming literature has shown that these representations can be activated without the individual knowing about or intending it, and then impact the individual’s evaluations,
emotions, and behaviors. A variety of environmental stimuli have been employed as activation triggers in these studies: not only verbal stimuli semantically related to the goal (as in many studies), but also material objects such as backpacks and briefcases to prime cooperation and competitiveness, respectively (Kay, Wheeler, Bargh, & Ross, 2004); candy bars to prime tempting hedonic goals (Fisbach, Friedman, & Kruglanski, 2003); dollar bills to prime greed (Vohs, Mead, & Goode, 2006); scents such as cleaning fluids to prime cleanliness goals (Holland, Hendriks, & Aarts, 2005); power-related features of the social context (Chen, Lee-Chai, & Bargh, 2002; Custers, Maas, Wildenbeest, & Aarts, 2008); and the names of close relationship partners (e.g., mother, friend) to prime the goals they have for the individual as well as those the individual characteristically pursues when with the significant other (Fitzsimons & Bargh, 2003; Shah, 2003).

Related evidence. In recent experimental work by Custers and Aarts (2005, 2007), a positive affective response was classically conditioned to the name of a particular goal. This manipulation increased the probability that the participant would pursue that goal over other possible alternatives, with the participant unaware of this influence on their goal pursuits. Extending this finding, Aarts, Custers, and Marien (2008) unconsciously manipulated both the goal of high performance (exertion) and the incentive attached to that goal; participants who had been subliminally primed with the goal of exertion did outperform a control group on the hand-grip squeezing task, but those primed simultaneously with both the exertion goal and positive stimuli outperformed everyone else. Moreover, Aarts, Custers, and Holland (2007) obtained the complementary effect: pairing negative stimuli with a goal increased the probability that the participant would disengage from the goal. Thus both the goal itself and the incentives associated with the goal can be manipulated unconsciously and will then influence task performance just as if the participant had been aware of the incentives and had consciously chosen the goal to pursue. These findings provide additional
support for the hypothesis that the same underlying brain regions and processes are mobilized in unconscious as in conscious goal pursuit, and that conscious goal pursuit makes use of preexisting unconscious motivational structures.

Emotional

Theoretical background. Emotions have long been taken to be the least controllable automatic activation event; Mowrer (1960) originally argued that emotions served to simulate the expected or anticipated outcomes of environmental events in a ‘safe’, internal manner so that behavior could be guided without having to actually experience those outcomes. For example, fear is an unpleasant, noxious emotional state that we seek to dissipate – but only internal, and not the greater unpleasantness of actual physical attack and damage that we might have stumbled into without the warning emotion. Accordingly, it would make a poor functional arrangement if our emotional states were instead under our own control; this would mitigate their important signalling function (see also Damasio, 1994). And if emotions were not automatically activated (Step 1) and difficult to prevent (Öhman & Mineka, 2001), we would not need to engage so often in the regulation of emotions after the fact (Gross, 1998).

Step 1. Emotional states can automatically trigger goals and motivational states of which the individual is not aware and does not consciously intend; anger, for example, induces approach motivation with associated feelings of greater efficacy and confidence; disgust induces a strong avoidance motivation towards all present stimuli; sadness triggers a motivation to change state (e.g., Haidt, 2001; Lerner & Keltner, 2000).

Step 2. Clues that this activation is unconscious and automatic include the participant’s lack of awareness of any carry-over influence of the prior emotional state on the subsequent task or decision, plus the finding that these emotion-induced goals produce outcomes that go against the
rational, best interests of the individual. For example, in the ‘endowment effect’ in behavioral economics research, people place a higher value on objects they possess than those they do not yet possess, resulting in the setting of higher prices to sell something they already have, compared to the price they’d be willing to pay to acquire the same object if they didn’t already own it. Selling an object for a higher price than one would pay for it makes perfect, rational economic sense.

However, Lerner, Small, and Loewenstein (2004) showed that the endowment effect can be moved around – even reversed – by manipulating the emotional states of participants. In an ostensibly unrelated first experiment, the researchers had participants think about something sad, or something disgusting, or neither in the control condition. Sad moods are known to trigger a goal to change one’s state. This goal was found to operate in the subsequent economic game, so that compared to the control condition, sad participants in the ‘buy’ condition were willing to spend more money (apparently to change state by acquiring the new object), and in the ‘sell’ condition, were willing to take less money (to change state by getting rid of the object) – thus reversing the usual endowment effect.

The disgust condition produced relevant findings as well. Disgust is an emotion known to trigger the goal of protection or rejecting objects from coming into contact with the body (e.g., Haidt, 2001). As would be expected if this protection motive had been unconsciously activated by the disgust prime, participants in the Lerner et al. (2004) study induced to recently feel disgust showed both lower buy (reducing chances of acquiring the object) and lower sell (increasing chances of expelling the object) prices, compared to the control condition.

The Unconscious as the Source of Behavioral Impulses

The present approach to the sources of human behavior is in harmony with contemporary theory and research in evolutionary biology, in which unconscious forces are understood to drive the
behavior of all living organisms. In the case of humans, these structures must have existed long
before the advent of conscious information processing capabilities. From these starting points the
hypothesis of unconscious behavioral guidance systems can be derived. Social cognition research
over the past quarter century has confirmed the existence of these unconscious guidance systems, in
that each variety of automatic process of relevance to social psychology – in the domains of
evaluation and attitude activation, social perception, and goal pursuit – has been found to be directly
connected to behavioral tendencies, without any need for conscious intention or awareness in the
production of these adaptive behaviors. Together, these findings support the present argument that
unconscious processes for adaptively guiding human behavior existed prior to the advent of
consciousness, and continue to generate behavioral tendencies today.

Such an idea is not new. Several theorists have suggested that the conscious mind is not the
source or origin of our behavior; rather, impulses to act are unconsciously activated, and the role of
consciousness is as gatekeeper and sense-maker after the fact (Eagleman, 2004; Gazzaniga, 1985;
James, 1890; Libet, 1986; Wegner, 2002). In this model, conscious processes kick in after a
behavioral impulse has occurred in the brain – that is, the impulse is first generated unconsciously,
and then consciousness claims (and experiences) it as its own. Take, for example, Libet’s (1986)
time of intention studies. In the Libet paradigm, participants are free to make a button-pressing or
other response whenever they choose (simulating the state of free will) and are asked only to note
when (by referring to a sweep-hand clock in front of them) they had made the intention to respond.
At the same time, the experimenter was measuring brain activation potentials associated with the
instigation of action. The finding, surprising at the time, was that the action potential consistently
came hundreds of milliseconds before the participant’s conscious awareness of intending to make
the response. Consistent with the present argument that our action impulses are generated for us
through unconscious mechanisms, the impulses, even in this paradigm emphasizing free will or action, came prior to the person’s conscious awareness of having made them.

Eagleman (2004) hypothesized from the Libet experiments that people do not generate intentions consciously but infer them based on perceptions of their own behavior. Recently, Banks and Isham (2009) tested this hypothesis in a new paradigm in which deceptive feedback was given to participants regarding the timing of their behavioral responses. When this feedback was slightly delayed (using an auditory beep representing when the participant began his or her behavioral response), participants’ estimates of when they initiated their action were also delayed. The estimate was thus an inference based on the perceived time of response initiation, and therefore participants could be ‘fooled’ by manipulations of that perception. These authors concluded from their experiments that “… the intuitive model of volition is overly simplistic – it assumes a causal model by which an intention is consciously generated and is the immediate cause of an action. Our results imply that the intuitive model has it backwards; generation of responses is largely unconscious, and we infer the moment of decision from the perceived moment of action” (our italics).

Wegner and colleagues (2002; Wegner & Wheatley, 1999) had already made this point in a different way, by showing how people’s feeling of having willed a given event to occur is an attribution or inference (not a direct readout of actual causation) based on key variables such as the timing of their thoughts of performing the action relative to the action occurring, through a novel paradigm in which these variables could be manipulated without the person’s knowledge. The right combination of these variables produced feelings in the participants of having willed the event when in fact it had not been under their control.

Thus many have proposed that consciousness is not the source of impulses to act. Yet to date there has been little said about where, exactly, those impulses do come from. Given the evidence
reviewed above, however, there now seems to be an answer to this question. There are a multitude of behavioral impulses generated at any given time from our unconsciously operating motives, preferences and their associated approach and avoidance behavioral tendencies, emotional reactions and their associated motivations, and mimicry and other behavior priming effects triggered by the mere perception of others’ behavior. There certainly seems to be no shortage of suggestions from our unconscious as to what to do in any given situation.

All of these separate types of input have their own direct connections to behavioral mechanisms, and they operate in parallel (see Bargh, 1997). And so there also must have been some mechanism to integrate the multiple parallel unconscious inputs into serial responses, because this is a problem we must have faced as a species in the distant past before the development of consciousness.

*How do multiple, parallel unconscious behavioral guidance systems interact?*

How do these parallel influences get channeled back through the bottleneck of having to act in real time? To presage the argument which follows, language is behavior (Clark, 1996): it is for doing things, for accomplishing goals with others (as opposed to by oneself), and so we might want to look to language production models – how ideas and thoughts are expressed – for insights as to how (other forms of) behavior are expressed. These linguistic production models have already tackled the problem of how parallel processes (thoughts, ideas, intentions) are transformed into serial speech acts (Dell, Burger, & Svec, 1997). It may be that serial, real time behavior *in general* follows the same principles, and even uses the same or a similar mechanism.

In other words, if language *is* action, then how language is produced may well be how behavior in general is produced. The relevant point about language production is that we do not usually formulate sentences in our mind prior to saying them. Rather, we may have some vague
ideas about what we want to say, the ideas or points we want to make, and these guide what we say, but the ways in which we express these intentions, out loud, with words, are opaque to us. That is, language is complex yet spontaneous, in most cases: it has a goal, but no pre-set concrete plan, yet it is nevertheless produced automatically and unconsciously (Bock & Levelt, 1994; Levelt, 1989).

Language probably made use of preexisting behavioral structures

An important clue in this regard is how quickly and suddenly, in terms of evolutionary time-scales, we acquired language (Pinker, 1994). It was not gradual, and did not depend on our brains growing to a certain critical size, for Neanderthal brains, which did not have language capabilities, were if anything larger than our contemporary brains (Calvin, 1989). Language is a complex skill that could not possibly be acquired so quickly in young children through normal, slow, trial-and-error learning processes (Chomsky, 1959); it develops spontaneously in nearly all children worldwide regardless of their levels of intelligence. The language production mechanism ‘takes a web of thoughts and outputs them in form of words spoken one at a time, without a conscious effort or formal instruction, and is deployed without awareness of its underlying logic’ (Pinker, 1994, pp. 101-102). In other words, it faced the same problem now posed to us in the case of behavior more generally: How to take multiple parallel behavioral suggestions and map them onto a world in which we can only do one thing at a time (see Bargh, 1997).

The speed with which we acquired language as a species, and the exponential advances in culture and knowledge we’ve made since then (see Diamond, 1992), suggests that as an ability it piggybacked or was ‘scaffolded’ (Williams, Huang, & Bargh, in press) onto an existing structure, or what Dennett (1995) called a ‘good trick’ – a solution that nature has come up with for a problem that tends to be used over and over again in nature (for example, eyes evolved independently in over 30 different species). The evolutionary theorist Calvin (1989) argued similarly that innate language
abilities themselves are quite recent, even rushed, additions to our genetic makeup, and as such as very likely exaptations of previously existing sequencing circuitry in the brain. What this means for present purposes is that not only did sophisticated unconscious modules evolve that give us today the building blocks of adaptive motives, preferences, and behavioral impulses, all operating unconsciously; there also evolved (indeed, *had* to evolve) a mechanism to integrate or interface these separate, parallel inputs into serial behavioral and judgmental responses. (Indeed, it has been proposed that the primary function of consciousness is to integrate the outputs of different action-oriented systems that are vying for skeletal muscle control [Morsella, 2005]). Our ability to take a vague thought and have it come out of our mouths in a complete coherent sentence, the production of which happens unconsciously, is a paramount example of this.

Miracles or mechanisms? Some concluding thoughts

We began this chapter by pointing out how unconscious processes are no longer considered as miraculous as they once were. But how conscious human cognition produces behavior might well seem miraculous as well given the standard cognitive-science approach to mental concepts – what Clark (1996, p. 55) termed the “product tradition”, or the *intellectual* aspect of concepts. In the product tradition, concepts are studied for what they contain, how they define the external objects, events, and qualities they stand for. Concepts, in this research tradition, are not for doing, they are for knowing. By contrast, the action tradition (typified by the work of Vygotsky and Goffman) has been concerned with how people use concepts (and language) to get things done in their daily lives; that is, the *experiential* aspect of concepts. The traditional cognitive science approach to concepts is all about their intellectual function and does not attempt to connect concepts to action – for example, in one recent *magnus opus* on concepts (Murphy, 2002), there is no mention of the development or nature of behavioral or action concepts—the term *action* does not even appear in the index. As
Lakoff and Johnson (1980/2003, p. 269) noted, “cognitive psychology is dominated by the old idea that concepts are all literal and disembodied.”

In contrast, for those who have been focusing on action itself for many years, and especially those of us who have focused on automatic or unconscious uses of activated concepts, they are all about the mediation of higher mental processes, including complex action and behavior in social settings. As showcased in the *Oxford Handbook of Human Action* (Morsella, Bargh, & Gollwitzer, 2009), these are exciting times because science is beginning to unravel the basic nuts and bolts of human action, the majority of which are unconscious. People are generally unaware of the sources of their behavioral impulses and of how their actions are successfully guided to completion; it is difficult indeed then to understand how conscious awareness can effectively guide action *without* massive support by unconscious guidance systems.

In the rest of the natural sciences, especially evolutionary biology and neuroscience, complex and highly intelligent design in living things is not assumed to be driven by conscious, intentional processes on the part of the plant or animal (e.g., Dawkins, 1976). As Dennett (1991, p. 251) put it, “in biology, we have learned to resist the temptation to explain *design in organisms* by positing a single great Intelligence that does all the work… We must build up the same resistance to the temptation to explain *action* as arising from the imperatives of an internal action-orderer who does too much of the work.” Especially, we would add, when there now exist such promising leads to how human behavior is generated and guided within the domain of unconscious processes.
References


