The Rationalizing Voter: Unconscious Thought in Political Information Processing

Milton Lodge
Distinguished University Professor
Department of Political Science
Laboratory for Political Behavior Research
Stony Brook University
Voice: 631-632-7663
Fax: 631-632-4116
Milton.Lodge@sunysb.edu

Charles S. Taber
Associate Professor
Department of Political Science
Laboratory for Political Behavior Research
Stony Brook University
Voice: 631-632-7659
Fax: 631-632-4116
Charles.Taber@sunysb.edu

Abstract: We review and evaluate a growing literature in social and political psychology on the ubiquity of unconscious thought processes and present a theoretical model, called *John Q. Public (JQP)*, which seeks to explain how citizens form and express their political beliefs, attitudes, and choices. Our most revolutionary claims are that people are generally unable to reliably report their political beliefs, attitudes, and behavioral intentions; that unconscious thought underlies all political deliberation, introducing important systematic biases, but paradoxically also providing the capacity for rational action in the face of severe cognitive limitations; and that conscious deliberation is typically more rationalizing than rational.

Under review at *Perspectives on Politics*.

Our intellect is not the most subtle, the most powerful, the most appropriate, instrument for revealing the truth. It is life that, little by little, example by example, permits us to see that what is most important to our heart, or to our mind, is learned not by reasoning but through other agencies. Then it is that the intellect, observing their superiority, abdicates its control to them upon reasoned grounds and agrees to become their collaborator and lackey.
— Marcel Proust

We are witnessing a revolution in thinking about thinking. Three decades of research in the cognitive sciences, backed by hundreds of well-crafted behavioral studies in social psychology and new evidence from the neurosciences, posit *affect-driven dual process models* of thinking and reasoning that directly challenge the way we political scientists interpret and measure the content, structure, and relationships among political beliefs and attitudes. Central to such models is the distinction between conscious and unconscious thinking, with hundreds of experiments documenting pervasive effects of unconscious thoughts and feelings on judgment, preferences, attitude change, and decision-making. What especially attracts our interest as political scientists to such dual-process models is the well-documented finding that unconscious thought processes are continuously at work, not only when people make snap judgments, but also when they think hard about important issues and decisions. These unconscious processes, moreover, constitute the overwhelming majority of human cognitive capacity with unacknowledged import for the character of political deliberation.

In this essay we review this literature with particular focus on four key postulates that directly underpin our model—called *John Q. Public (JQP)*—of how citizens form and express their political

beliefs, attitudes, and choices.[1] Perhaps our most revolutionary claim is that people are generally unable to reliably and veridically report their political beliefs, attitudes, and behavioral intentions, and their conscious deliberations are typically more rationalizing than rational.

Automaticity

What people think, feel, say, and do is a direct function of the information that is momentarily accessible from memory—be it the recall of facts and feelings, the recollection of experiences, or the turning of goals into action. Contemporary theory distinguishes between conscious and unconscious (explicit/implicit, deliberative/automatic) processing in the formation and expression of beliefs, attitudes, goals, and behavior. Implicit attitudes are defined "as introspectively unidentified (or inaccurately identified) traces of past experience that mediate favorable or unfavorable feelings toward an object."[2] For example, the feelings of pride and in-group solidarity that arise when flags wave or patriotic music plays in the background of political events, the subtle confidence felt in the presence of tall political candidates or infatuation for attractive or charismatic ones, and the unease experienced by some voters at the prospect of African-American or women leaders all influence political thinking outside conscious awareness.

The research on automaticity demonstrates that beliefs, feelings, and behavioral intentions will, if "contiguously activated" become so strongly connected in memory as to become *unitized* in a network of interdependent associations that enter the decision stream spontaneously on mere exposure to a "triggering event." Automatic processes, in which thoughts, feelings, and intentions come to mind unconsciously, on a time scale of milliseconds, contrast with the more demanding processes people engage in when they have sufficient time, motivation, awareness, and cognitive

[1] For a formal, computational model of *JQP* see Kim 2005; Kim, Taber and Lodge 2007.

[2] Greenwald and Banaji 1995, 8; for a general review, see Bargh 1997.

resources to deliberate. Process matters: with the repeated association of thoughts to feelings, beliefs become affectively charged; feelings motivate intentions, and these plans direct behavior. From this perspective, Antonio Damasio is right in seeing the brain as a "thinking machine for feeling," and William James right in saying "thinking is for doing."[3]

Realizing that the human mind processes information both consciously and unconsciously, theorists have proposed a conceptual distinction between attitudes that are the products of introspection and those that occur implicitly, outside of conscious appraisal. The labeling of one mode of processing as "conscious" emphasizes the reflective, deliberative character of one's responses to an "object"—whether person, place, event, thing, or idea—which generally (but not necessarily) involves verbal reasoning. Deliberative processes are cognitively effortful, demanding of attention, time consuming, and presumed to be based on an intentional memory search for relevant facts and considerations. Conversely, automatic processes—whether the immediate activation of cognitive associations (e.g., Bush is a Republican), the spontaneous activation of feelings (Republicans are evil; Democrats are dumb), or those habitual actions that operate "mindlessly"—are involuntary, fast, immediate, top of the head, and unlike conscious processes can be activated *even* when the individual's conscious attention is focused elsewhere. In sum, people are frequently unaware of the specific situational and contextual factors (call them "primes") that bring to mind the thoughts, feelings, and intentions that appear introspectively to be the outcome of a deliberative evaluation of the evidence. Implicit processes, moreover, have been found to produce sound decisions, oftentimes better than those based on careful deliberation.[4]

Priming effects—whether consciously recognized or not—are ubiquitous in everyday life. The men and women of beer or car commercials are always unusually attractive; the smokers in cigarette ads look preternaturally healthy; the men touting erectile dys-

[3] Damasio 1994a; and James 1890.

[4] See for example Forgas 2000; and Wilson 2002.

function medications appear too virile to need them. Such "incidental," normatively-irrelevant primes are well-documented in the political domain as well as in the world of commercial advertising, as when voters' evaluations are influenced by the attractiveness of the candidate, by the amount of time ABC anchorman Peter Jennings smiled when reporting on Reagan over Mondale in the 1984 presidential race, or by the racial cues in the Willie Horton ad attacking Michael Dukakis in 1988.[5] These effects, like the impact of a sunny day on survey respondents' evaluations of their life satisfaction, are strongest when the observer is unaware of the prime's influence.[6] If in fact the contingency is made explicit the biasing effect may be wiped out, or ironically, overcompensated for.

While priming paradigms are used primarily in experimental settings to tease out automatic influences on cognition, given the obvious relevance to persuasion it was to be expected that these laboratory methods would find their way into advertising as hidden persuaders and then quickly into the selling of the president. In the 2000 presidential election campaign the Republican National Committee aired a TV ad nationwide 4,400 times, costing $2,576,000, attacking Gore's prescription drug plan. When the final segment of the ad is slowed down, the word "RATS" flickers nearly imperceptibly across the screen. Shown at regular speed "RATS" appears at the near-subliminal speed of 300 ms, superimposed over the words "Bureaucrats Decide." The ad's creator said it was not his intention to create a subliminal ad, but rather to make the ad more visually interesting by flashing part of the word "bureaucrats" on the screen. "It was," he said, "just a coincidence" that the letters popping out of "bureaucrats" spelled out the negative prime "rats". Such denials notwithstanding, experimental tests of the "rats ad" show a significant negative impact on evaluations of Gore and his drug plan.[7] Both inside the lab and in

[5] Sigelman 1990; Mullen, Futrell, Stairs, Tice, Baumeister, Dawson, Riordan, Radloff, Goethals, Kennedy and Rosenfeld 1986; and Mendelberg 2001.
[6] Schwarz and Clore 1983.
[7] For example, Weinberger and Westen 2007.

the real world, priming effects are proving to be critically influential in how information is encoded, retrieved, interpreted, evaluated, and acted upon.

Hot Cognition

The second postulate driving *JQP* and all affect-driven dual-process models is the "*hot cognition hypothesis,*" the claim that with repeated co-activation socio-political concepts become positively or negatively charged and this affective charge becomes directly linked to the concepts in long-term memory. Thereupon "feelings" come to mind spontaneously upon mere activation of the object represented in long-term memory, be it a person, group, issue**,** event, symbol, or idea. By election eve, for example, most citizens will have formed impressions of all major candidates, parties, and issues, and these feelings will be inescapable when deliberating about their election choices, especially so for those citizens who have given more thought to the election.[8]

The Primacy of Affect

It is now well-documented that feelings enter the decision stream *before* any cognitive considerations come consciously to mind. Neurological studies of both mice and men suggest that the "affect system" follows a "quick and dirty" pathway that prepares organisms for approach-avoidance behavioral responses moments before any conscious appraisal. The behavioral and neurological evidence is clear that affective responses are activated immediately on mere exposure to a triggering event, even when as we will see, exposure to the stimulus object occurs outside of conscious awareness. Given the "primacy of affect" all thinking, reasoning, and action—whether conscious or unconscious—is embodied by a felt sense of like or dislike, which is experienced through what Antonio Damasio calls a "somatic marker," a gut level sense of

[8] Lodge and Taber 2005.

something being good or bad. In direct contrast to much of Western thought which treats feelings as befuddling rational decision-making, Damasio's somatic marker hypothesis posits that bodily feelings normally accompany our mental representations of the anticipated consequences of real or imagined decisions, with thoughts uncoupled from feelings only in pathological cases. Directly linked to one's history of rewards and punishments, somatic markers serve to facilitate responses that were successful in the past. Affect then anchors all judgments and is always a central and oftentimes the single most dominant piece of information guiding judgment and choice. Indeed, feelings toward the candidates on election eve may be the most powerful and useful information available to citizens, making truly dispassionate deliberation neither possible nor desirable.[9]

On-line Processing

But how are these feelings updated and in response to what? The on-line (OL) model holds that beliefs and attitudes are constructed in real time, at the moment of comprehension, from the momentarily activated associations in memory, when an object is before your eyes, so to speak. When people form or revise their overall impressions of persons, places, events, or issues they spontaneously extract the affective value of the message and, within milliseconds, update their summary evaluations of the object. These "running tallies"—representing an integration of all prior evaluations of the object—are then restored to memory where they are readily available for subsequent evaluations. From this OL constructionist perspective, affect infuses the encoding, retrieval, and comprehension of information, its expression as a preference, and readies us to act aversively or appetitively in accord with our feelings.[10]

The affective tally associated with social concepts inevitably colors all phases of the evaluation process, sometimes explicitly,

[9] Damasio 1999; LeDoux 1996.

[10] Hastie and Park 1986; Lodge, Steenbergen and Brau 1995.

other times not, sometimes for good, other times not. If these transitory effects were the full extent of the influence of on-line tallies on evaluations they would nonetheless be important in that momentarily activated affect is integrated in real time into OL tallies and thereafter anchors subsequent evaluations. But we now have good reason to believe that on-line affective processing comes into play, perhaps even more strongly, when one consciously and systematically weighs the pros and cons.

Automatic effects on memory and judgment are ubiquitous in our everyday lives. Virtually all mental representations appear to be "prime-able"—activated incidentally or unobtrusively in one context to influence one's thoughts, feelings, goals, and even complex behaviors in another context—without the person necessarily being aware of having been influenced. Priming effects, moreover, have been demonstrated on virtually all higher mental processes.[11]

These effects—the unconscious linking of feelings to thoughts to goals to behavior—require a revolution in how we think about and model citizens' mental representations of the world and the processes involved in the formation and expression of their political beliefs, attitudes, and behavior. To the extent that this dual-process model provides a valid account of the structures and processes of mental representations, much of our everyday life operates automatically. And, given that feelings enter the decision stream before cognitive considerations become conscious, we believe that the way we political scientists conventionally model the relationship between beliefs and attitudes is misdirected—for most people most of the time the casual arrow flies from feeling to thinking, from affect to cognition.

This the case, we expect that most citizens most of the time will be *biased reasoners* who find it near impossible to evaluate new, attitude-relevant information in an evenhanded way. When exposed to challenging, attitudinally-incongruent information people routinely rationalize the facts, figures, and arguments that they cannot effortlessly discount, depreciate, denigrate, or deny. Like

[11] For examples, see Bargh 1997.

the Bush 43 administration, citizens are prone in their everyday lives to make facts fit their priors. Even when held accountable for their judgments, people are prone to interpret attitudinally congruent evidence as inherently stronger than attitudinally incongruent evidence; spend more time and exert more energy counterarguing facts and arguments that challenge their priors; and guard against contrary information by selectively seeking congruent information. As a consequence of this motivated search, seizure of supporting evidence, and closure before all the evidence is in, attitudes are prone to become more extreme in the face of a balanced set of pro and con arguments, with all these biasing effects strongest for the most knowledgeable and committed citizens, that is, for people like you and me. Yes, of course, under the right circumstances and given enough motivation to be prudent we can confront challenging evidence and correct for biases, but correction processes are very effortful and no guarantee of success.[12]

In this essay we set forth our dual-process model (*JQP*) of the architecture and mechanisms that can account for when, how, and why thoughts, feelings, goals, and behavioral expectations come to mind automatically to promote behavior. In contrast to the minimalist assumptions of utility maximization in rational choice theory, we propose a psychologically realistic account of how citizens think, reason, and act in accord with their feelings toward political leaders, groups, and issues. We take a constructionist approach whereby the content of one's thoughts and the coloration of feelings change moment-by-moment in response to unnoticed "priming" events that link changes in the environment to changes in political beliefs, attitudes, and behaviors.

The *John Q. Public* Model of Automatic Political Beliefs, Attitudes, and Behaviors

A cornerstone of any psychological model of political reasoning is the citizen's mental representation of the world—their factual

[12] Kunda 1990; Taber and Lodge 2006; Wilson 2002.

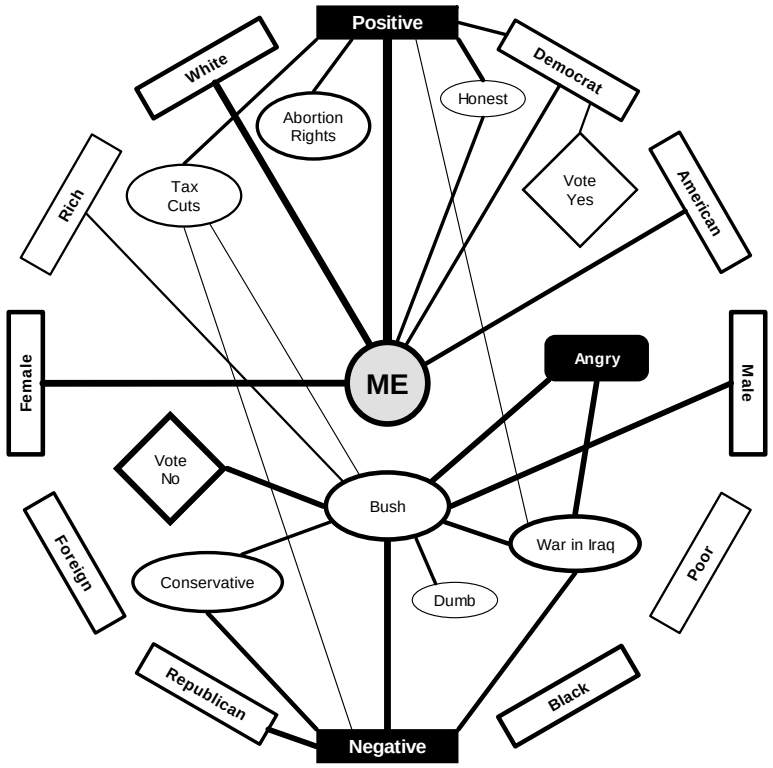
knowledge and predilections. These recorded experiences, functionally speaking, require a vast *long-term memory* (LTM) for storing facts, beliefs, feelings, habits, predilections, and behavioral predispositions, plus a *mechanism* for "moving" such political objects as leaders, groups, events, and issues from LTM into *working memory* (WM) where they can be attended to. Conscious attention is very limited, roughly to the magic number 7 ± 2 bits or chunks of information, hence the need for heuristics, habits, and other simplifying mechanisms for thinking and reasoning.

Figure 1 illustrates the associative architecture of one hypothetical woman's political knowledge structure, represented as configurations of *nodes* linked one
to another in a network of associations. Beliefs, attitudes, emotions, and intentions are represented as links among basic memory objects. For example, Figure 1 depicts the beliefs "I am a Democrat" and "Bush supports the war in Iraq," the attitudes "I don't like Bush" and "I support abortion rights," the emotion "the war in Iraq makes me angry," and the behavioral intention to "vote Democrat." Were we to tap a citizen's full political knowledge structure, there might be tens of thousands of conceptual nodes, with a complex network of links to associated concepts (perhaps George W. Bush's perceived stands on issues or traits, affective responses, and maybe an inferential abstraction or two—e.g., that he is conservative). The strength of links varies from weak to strong, a function of the number of times and the context within which the concepts have been co-activated. Hence, memory objects vary in their *accessibility*—the ease with which a stored object lying dormant in LTM can be retrieved to influence the decision stream—which is a function of the node's strength and its connections to other concepts.

But how does information move into consciousness? Spreading activation provides the mechanism. An object node in LTM switches from dormancy to a state of readiness with enhanced potential to be moved into conscious WM, either by direct exposure

to the object or because it is closely linked to an object of thought. The rise time from dormancy to activation threshold is almost instantaneous (100-200 milliseconds). Were you thinking of George

Figure 1: The Associative Structure of
Political Beliefs, Attitudes, and Intentions



Washington's false teeth? Are you now? Activation also decays rapidly so that a given node will drop back to its baseline level of potentiation in about a second if there is no further source of activation. Imagine a person reading "President Bush" in a newspaper headline. Without perceptible effort, the concept BUSH (G.W.,

not a shrub) becomes activated and pops into consciousness. Even more important for our purposes, activation spreads along the network of links to related concepts, as in Figure 1, where BUSH primes the strong semantic association to CONSERVATIVE as well as to associated beliefs ("he may cut taxes"), traits (he's none too smart), feelings ("he makes me mad"), and behavioral intentions ("I plan to vote against him").

It is useful to think of spreading activation as producing preconscious expectations. Consider again the activation of the concept BUSH from a newspaper headline. If an association (his Republican label or his stand on gun control) is energized, it takes substantially less processing time and effort to become activated and thereby has a better chance of getting into consciousness, of being processed faster, and consequently "framing" the perception, recognition, and interpretation of subsequent information. Conversely, spreading activation can *inhibit* the processing of unexpected categories. When a concept is encountered unexpectedly, more bottom-up processing is necessary before it passes threshold to enter consciousness. If the word "Einstein" were processed initially, this would certainly inhibit the recognition of semantically unrelated concepts (such as BUSH).

Central to affect-driven dual process models of attitude is the *hot cognition postulate*, which brings feelings center-stage in human information processing in claiming that all socio-political concepts are affect-laden. As depicted in Figure 1 attitudes appear as links between concepts in memory and positive and/or negative affect. All political leaders, groups, issues, symbols, and ideas thought about and evaluated in the past have become affectively tagged—positively, negatively, or both—and with repeated co-activation an evaluative charge is linked directly to the concept in long-term memory. This affective tag represents the value of social "objects" as good, bad, or ambivalent.

With repeated evaluations an evaluative tag comes inescapably to mind upon mere exposure to the associated object—"what is wired together fires together"—thereby signaling its affective coloration. At the moment one recognized the image on the TV as

the President, one's feelings about "G.W." Bush come to mind, followed shortly thereafter by his strongest cognitive associations. These accumulating positive and/or negative affective charges stimulate physiological changes within the body which are experienced via a gut level sense of good or bad, which if strong enough to call for an answer to the question "why do I feel this way?" may be experienced as a discrete emotion (e.g., anger, fear, joy).

The impact of context on evaluations follows directly from the model. If "jobs" is primed for a working class citizen, "business" may be seen in a positive light, while in the context of "greed" "business" will likely be evaluated negatively. Note too that one is ambivalent when there are links to both positivity and negativity, as with "tax cuts" and "the war in Iraq" in Figure 1. From this constructionist perspective, the evaluation of an object represents the integration of multiple sources of affective information, from the object itself (Bush is negative) as well as from its strongest associations (Republican is negative).

Simple though it be, the direct linking of feelings to concepts, to goals, and to behavioral intentions has profound implications for human information processing. The associative strength between an object (e.g., politician) and its evaluation is conceived as varying along a continuum from nil—an object with little or no affective association—to a "crystallized attitude," that is, an object with a strong, chronically-accessible, evaluation. Whereas weak or non-attitudes require effortful, piecemeal, bottom-up construction, the stronger the association between an object in memory and its affective evaluation the less time and effort needed to bring the attitude to mind, with objects carrying strong affective links activated spontaneously on their mere exposure, without the observer necessarily being aware of having perceived the triggering event.[13]

Hot cognition helps solve the problem posed by the French scholastic Buridan, who said: "If a hungry ass were placed exactly between two hay-stacks in every way equal, it would starve to death, because there would be no motive why it should go to one

[13] Bargh, Chaiken, Govender and Pratto 1992.

rather than to the other."[14] Most humans, unlike Buridan's ass, are equipped to solve such "equilibrium problems" by tagging the valence of goals which thereupon facilitates the making of quick, intuitive choices. That affect permeates the entire decision-making system ensures that beliefs, feelings and actions will typically cohere. When things go wrong, of course, there is a good chance that both thoughts and feelings will conspire to promote a misguided response.[15]

This constructionist perspective implies that the evaluations a citizen might report in an opinion poll or vote choice reflect the integration of feelings associated with one's history of conscious and unconscious political evaluations. Immediately and without intentional control, a perceived candidate, issue, group, or idea is classified as either good or bad, and in a matter of milliseconds this evaluation facilitates a behavioral disposition toward the stimulus.[16]

The Updating of Evaluations

The integration of affective tags into global evaluations of persons, groups, ideas, or events appears to be governed by an anchoring and adjustment heuristic which weighs early information more heavily than information experienced later in the decision stream. This powerful effect of primacy on impression was first suggested by Solomon Asch, who argued that the very same descriptors of a person would produce a different holistic impression, depending on their order of presentation.[17]

Try the following thought experiment: On a scale that ranges from "highly favorable" (+4) to "highly unfavorable" (-4) evaluate person A, who is described as:

[14] Quoted in Brewer's *Dictionary of Phrase and Fable*. Henry Altemus, Philadelphia, 1898; available at www. Bartleby.com.
[15] Thagard 2006; Wilson 2002.
[16] Morris, Squires, Taber and Lodge 2003.
[17] Asch 1946.

faultfinding, awkward, cool, sentimental, athletic, and smart.

Count backwards by 7s from 100 until you reach 65. Now evaluate person B, who is described as:

smart, athletic sentimental, cool, awkward, and faultfinding.

Experimental subjects rated person A as "slightly unfavorable" (-0.7) and person B as "moderately favorable" (+1.4), despite the fact that only the order of the trait descriptors changed.[18]

There is now a great deal of evidence that impressions of people form spontaneously and quickly, anchor on early information, and adjust insufficiently to later information. The formation of evaluations is dominated by first impressions.[19] By contrast, the recall of cognitive considerations is most strongly influenced by recent information. Information that comes early in a political campaign is likely to strongly influence feelings toward candidates, even though later information may be better remembered.[20] It does not escape our notice that this pattern of primacy effects on impression formation but recency effects on recollections renders recall-based measures suspect as indicators of why people favor one person or idea over another.

The Primacy of Affect

Contemporary models assert the *primacy of affect* in positing that feelings enter into the decision stream faster and earlier (both in cognitive and evolutionary time) than do semantic associations. People feel their opposition to the Iraq war before any facts about the war come to mind, and these feelings influence what cognitive considerations come to mind. Over time, this affective mechanism helps to structure political knowledge, though by structure we have in mind associated connections between concepts rather

[18] Anderson and Barrios 1961.
[19] Uleman and Bargh 1989.
[20] Lodge, Mcgraw and Stroh 1989; Zaller and Feldman 1992.

than ideological constraint. Neurophysiological evidence suggests that the "affect systems" form "quick and dirty" pathways in the brain to facilitate approach-avoidance attitudes and behavioral responses. It appears that affect and cognition are separate, yet interdependent systems, only separable in pathological cases.[21] What is clear on the basis of our own experimental evidence and a great deal of evidence from other fields is that affective responses come to mind spontaneously within 100-250 ms of exposure to a social stimulus, and appear to enter the evaluation process unconsciously moments before cognitive considerations come to mind.[22]

Even when the conscious mind is focused elsewhere, automatic evaluative processes prepare the individual to make an affectively congruent response. As a preview of more to come, consider two experimental demonstrations of the linkage between feelings and approach-avoidance behavior:

In a subtle but remarkably direct demonstration of this link between feelings and behavior, social psychologists Chen and Bargh instructed half their subjects to pull a joystick toward themselves when positive words appeared on the computer screen and push the lever away for negative words, while the other half received the opposite push-pull instructions. Results confirmed that subjects were faster to pull the joystick toward them for pleasant words and push away for unpleasant words, a result we see as telltale evidence for the central role played by simple affect in triggering basic approach-avoidance behaviors.[23]

A more complex behavior: In 1997, a group of cognitive neuroscientists at the University of Iowa published an influential paper in *Science* that vividly captured the importance of automatic feelings for decision making. The Iowa group described a simple gambling game in which participants chose cards, one at a time, from four decks, two red and two blue, with each card carrying a

[21] Damasio 1994b; LeDoux 1996; Zajonc 1980.
[22] Burdein, Lodge and Taber 2006; Crites, Cacioppo, Gardner and Berntson 1995; Lodge and Taber 2000; Lodge and Taber 2005; Morris, Squires, Taber and Lodge 2003.
[23] Chen and Bargh 1999.

monetary payoff or loss. Some red cards provided large payoffs, others carried large losses, and the expected value of the red decks was negative. By contrast, the blue decks were contrived to provide moderate payoffs, with occasional small losses, that yielded a positive expected value. Players in the game, like those facing uncertain prospects in nature, had to evaluate their options and select strategies through trial and sometimes painful error. The quicker they determined which fork in the road presented danger and which opportunity, the quicker they realized benefits and avoided losses.[24]

Players in this game, it turns out, are able to articulate the underlying structure of the game within about 80 cards, and they can typically report a hunch about the red decks after sampling 50 cards. This corresponds to the "deliberative" solution to the game, and if this were the only mechanism through which players could intuit the game, there would have been quite substantial losses. But this was clearly not the only mechanism, *nor even the one that actually drove behavior*, since these same subjects displayed behavioral adaptations (i.e., they avoided red decks) long before they arrived at their first conscious glimmer of a hunch, typically after about 10 cards! They showed, moreover, physiological symptoms of stress (e.g., sweaty palms) in response to red cards at precisely the same time that their behavior began to change, and again, long before they were consciously aware of the "red danger." Taken together with the "joystick" experiment and dozens of other empirical examples, this remarkable demonstration of a somatic marker illustrates quite directly the adaptive benefits of automatic processing for organisms that face dangerous and uncertain choices.

In direct contrast to much of Western thought, which treats affect, feelings, and emotion as irrational interference that befuddles decision-making, *JQP* follows the lead of this recent neuro and social psychological evidence, in connecting positive and negative feelings aroused by external events and internal thoughts to goals, choices, and behavior via Damasio's somatic marker hy-

[24] Bechara, Damasio, Tranel and Damasio 1997.

pothesis.[25] Directly linked to rewards and punishments previously experienced in social situations, these gut-level feelings automatically "signal" whether a situation, event, or option is good or bad, potentially threatening or rewarding. This bodily sensation may be felt below conscious threshold as an intuition, or in other cases it may be experienced as intense arousal, demanding cognitive appraisal and emotional identification. However it is experienced, a somatic marker ensures that options will be viscerally embodied, thereby facilitating approach or avoidance behaviors by signaling the prospect of pleasure or pain. Somatic markers allow the brain to use affect as information and promote quick, efficient, spontaneous responses to what should be approached and what avoided. Ironically, the very same processes that so often promote bias in human thought are also responsible for rational action.

The affective labeling of options helps prevent decision calculations from becoming so complex and cumbersome that decisions would be impossible. Indeed, hot cognition is what makes decision-making possible. Feelings provide feedback about the unconscious processes that precede conscious consideration. Because brain processing capacity is greater and faster than conscious appraisals, this affect heuristic first provides an alternative to and later perhaps a complement to conscious processing. Without the weighting of goals by affect none would be more important than another and we would consequently end up like Buridan's ass, unable to choose among preferences.

Experimental Tests of the Automaticity of Political Beliefs and Attitudes

To call a process "automatic" it must satisfy four criteria: it must be *spontaneous*, *unconscious*, *uncontrollable*, and it must *expend little or no cognitive resources*. What is important here is that in many familiar situations (as well as in such uncommon settings as a survey interview) automatic processes will directly impact the expression of

[25] See also Marcus, Neuman and MacKuen 2000.

subsequent evaluations, judgments, goals, decisions, and actions with little or no conscious or deliberative guidance. Bargh draws an important distinction between preconscious and postconscious automaticity. In postconscious automaticity one is aware of the stimulus but not of its influence on thought, feeling, or behavior, whereas in preconscious automaticity the priming event occurs below the threshold of conscious awareness so that the observer is not subjectively aware of having been exposed to the priming stimulus. But who cares what happens in the blink of an eye?[26]

Postconscious Automaticity

Telltale evidence of postconscious processing is routinely discerned in public opinion surveys, showing up as question-wording and question-order effects. For example, a *Washington Post* opinion poll asked a national sample of Americans in November of 2002, when President Bush's approval rating was in the mid 60s, whether the country was headed "in the right direction" or "was seriously off in the wrong direction." Immediately before or after this question, respondents were asked whether they approved or disapproved of the job Bush was doing as President. A postconscious "Bush effect" is implied by the finding that 42% of those asked the Bush question first believed the country was headed in the right direction, whereas only 34% felt that way when the Bush question was asked second.

Our model predicts that priming effects—whether sparked by a President's name, upbeat music in the background of a commercial, the sound of prison doors slamming shut, or even having "rats" jump out of "bureaucrats"—would produce similar biasing effects on information processing. As with flags and other symbols in the backdrop of presidential speeches, the more subtle and unobtrusive the "manipulation" the stronger the effect should be.

The Implicit Association Test (IAT) is currently the most popular procedure for measuring postconscious automaticity. Studies using the IAT find implicit stereotyping behaviors that

[26] Bargh 1997.

often diverge from explicit (i.e., consciously processed) measures such as the Modern Racism Scale.[27] Implicit measures, being outside direct control, are more resistant to social desirability and impression management biases. That is, people cannot consciously censor their responses to appear politically or socially "correct" as they can for most explicit measurement strategies. Even more important, implicit measures reveal thoughts and feelings that are not introspectively available to the respondent. The uncomfortable fact is that people often do not know their own minds, and even one's sincere explanation for preferring one thing over another may be more rationalization than rational.

A compelling test of postconscious automaticity in on-line processing was carried out by Tilman Betsch and his colleagues. They had their subjects watch a series of 30 videotaped TV commercials, which they were told they would have to later recall and evaluate. Simultaneously, these subjects performed a second, cognitively-demanding distractor task: they were asked to read aloud the changing stock prices of five hypothetical companies presented on a crawler at the bottom of the TV screen. Though participants were led to believe that their ability to remember and evaluate the TV commercials under pressure of an irrelevant distraction was the primary purpose of the study (recall of the commercials proved to be very good), the researchers were actually interested in how the viewers would track the stock ticker. In a surprise test, subjects were asked their preferences for the five companies. As predicted by the on-line processing model, participants were unable to recall the pertinent stock information, yet

[27] Greenwald and Banaji 1995; see demonstrations of the IAT at http://implicity.harvard.edu. Nosek and Smyth 2007 conducted a meta-analysis of studies that used the IAT to compare explicit and implicit attitudinal responses across 57 attitude domains—including such concepts as Republican and Democrat, straight and gay, creationism and evolution, fat and thin, black and white—finding implicit by explicit correlations corrected for measurement error ranging from 0.18 to 0.68. Latent variable structural models on these data demonstrated that a two-factor model, including implicit and explicit attitudes as separate factors, was superior to a single-factor specification, strongly suggesting that implicit and explicit attitudes are separate constructs.

their summary, rank-ordered evaluations correlated positively and strongly with the actual performance of the 5 companies. This result points to the automaticity of on-line evaluations: subjects accurately evaluated the companies' stock performances even when their attention was actively focused on an unrelated, attention-demanding task.[28]

Current research is going beyond the automaticity of evaluation processes to focus on the postconscious activation of complex social behaviors. A now classic experiment primed the concept "elderly" by asking subjects to unscramble sentences that included such words as Florida, worried, old, lonely, gray, and bingo. The behavioral dependent variable was the time (measured in seconds) it took subjects after the priming experiment to leave the lab and walk to the elevator. Those primed by the concept elderly took significantly longer than control subjects to walk the 30 meters to the elevator, even though none of the primes for elderly referenced slowness of gait and the study participants were college students, not old folks. Yet their mental representations of the elderly activated a rich behavioral script that included slow walking. Another study (among dozens of similar demonstrations) primed one group of subjects to the concept "professor" and another to "soccer hooligan" and found (sigh of relief) that the professor-primed group correctly answered more *Trivial Pursuit* questions than did those exposed to the hooligan primes.[29]

A series of remarkably subtle experiments which suggests the ubiquity of postconscious priming in everyday life investigated the effects of simple business primes (e.g., pictures of boardroom tables, men's and women's business suits, attaché cases) on competitive behavior. The basic premise is that these common objects carry implicit psychological meaning (business is competitive) and would, even if presented unobtrusively, facilitate competitiveness in experimental subjects. Across five studies the design was to first engage participants in a postconscious business-related priming task, and then in an ostensibly unrelated second study present

[28] Betsch, Plessner, Schwieren and Gutig 2001.

[29] Bargh, Chen and Burrows 1996; Dijksterhuis and van Knippenberg 1998.

them with one or another behavioral task in which participants could act cooperatively or competitively.[30]

Study 1 asked treatment subjects to match business-related pictures to word labels, while control subjects performed the same priming task for such non-business objects as kite, sheet music, and toothbrush. All participants were then asked to complete 24 word fragments, 9 of which connoted competition, among them (w)in, (p)ower, wa(r), and one fragment, c__p___tive. While none of the participants reported awareness of the relevance of the priming manipulation, those primed with business objects completed significantly more competitive word fragments than the control group. Moreover 24 of 34 treatment subjects saw "competitive" in the fragment c__p___tive, while only 13 of 33 in the control group did so.

But can people be postconsciously cued to actually behave more competitively? A second study addressed this question by first priming subjects with business or non-business pictures and then asking them to play an Ultimatum Game, in which one player chooses how much of $10 to offer to another player in a one-time, take-it-or-leave-it proposition. Here too the results show strong priming effects, even though the participants were consciously unaware of any connection between the picture primes and their subsequent behavior. All but one of the control subjects offered an even split, but 7 of the 11 who had been primed with business images offered significantly less.

The final in this series of studies saves the best for last. The Ultimatum Game again, with all subjects in the role of choosing how much of $10 to offer another (unseen) player, but now there is no picture-priming task. Instead, subjects write down their offer in one of two settings: Half made their offer in a room with a long wooden conference table on which lay at the far end a leather briefcase and before them a black leather portfolio and silver, wide-barrel, executive style pen. The experimenter removed the ultimatum offer sheet from the briefcase and instructed them to write down their offer and replace the sheet in the briefcase. For

[30] Kay, Wheeler and Bargh 2004.

the other half of the participants—same room—a student's backpack replaced the briefcase at the far end of the table, a cardboard box substituted for the executive portfolio, and the take-it-or-leave-it bid was made using a wooden pencil. After making their offer, all were asked to list the factors that contributed to their offer. None indicated being influenced by any of the objects in the room, and yet the results show significant effects of condition on offers, with all 10 subjects in the backpack condition opting for a 50-50 split, while only 6 of 12 in the business setting did so.

These studies and many more demonstrate the influence of unnoticed primes on perceptions, social judgments, and behavior. Of special note here is that in every case, study participants were consciously aware of the environmental primes, but were unaware of their biasing effects. Similar effects, we expect, are ubiquitous in everyday life outside the laboratory, including the development and public expression of political beliefs and attitudes.

Preconscious Automaticity

Preconscious automatic responses—whether feelings, thoughts, motivations, or overt behaviors—occur spontaneously, within milliseconds of a triggering event, without conscious attention, awareness, intention, or monitoring. To illustrate, let us describe a trio of experiments that usher in themes that we will focus on when describing our own studies of automaticity in the evaluation of political leaders, groups, and issues: (1) preconscious processing, (2) the automaticity of affect, and (3) the primacy of affect.

Psychologists Dijksterhuis and Aarts set out to test the hypothesis, rooted in evolutionary theories of automatic vigilance, that people process negative stimuli more quickly than positive stimuli. Other research has shown that negative events and objects command more attention than positive, but these studies look one step earlier in the process in asking whether negative stimuli are detected faster and easier at the preconscious level.[31]

[31] Dijksterhuis and Aarts 2003.

Their Study 1 asked whether participants would be able to detect positive or negative words flashed on a screen at the subliminal speed of 13.3 ms, far too fast for conscious recognition. For half of the trials, a positive or negative word appeared, for half of the trials no word appeared. The subjects, fully informed in advance about the expected frequency of words but not of their valence, were asked at each trial whether they thought a word had been presented. Not surprisingly at this subliminal exposure time, none of the subjects could consciously discriminate whether a word or nonword flashed on the screen, yet they correctly guessed significantly more negative than positive words. In short, what participants reported to be pure guesswork turned out to be systematically biased in favor of detecting negative stimuli.

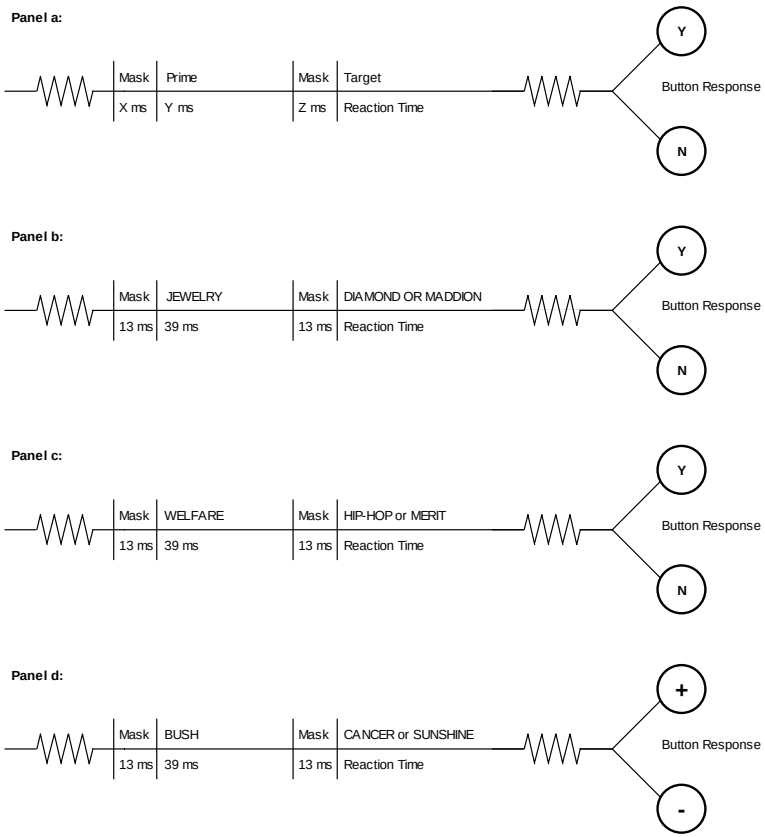
But at what level did they perceive this negativity? Study 1 showed that negative words were detected faster, but it did not ask whether subjects preconsciously felt the valence of the words. Studies 2 and 3 take this next step, asking participants to press one key for positive words and another for negative. Again, words were presented at 13.3 ms, but now either a positive or negative word was flashed on every trial. Results confirmed predictions: the proportion of correctly identified negative words was significantly higher than correctly identified positive words, despite the fact that again participants believed they were guessing randomly. But can we yet be sure that this preconscious vigilance for negative stimuli was truly affective? Perhaps the semantic meaning of negative words is somehow easier to process.

Study 3 eliminated this possibility by, in addition to the evaluative task described above, asking participants to guess which of two similarly valenced words, presented explicitly on the screen, was a synonym of the subliminally presented word. The results were striking: Although subjects detected significantly more negative than positive words, they were not able to reliably identify any synonyms, negative or positive—that is, they were unable to discriminate the meanings of the concepts. This then is a clear demonstration of preconscious affective processing and telltale evidence of a disjuncture between affective and semantic processing,

with people able to "feel" that something is good or bad even though they are unable to tell you what it was they saw.

Two experimental paradigms have proven particularly useful in demonstrating preconscious influence on higher order mental processes. One—*the lexical decision task*—is used extensively in cog-

Figure 2: Lexical and Attitude Priming Paradigms



nitive psychology for determining the meanings of concepts and their associations, the other—*the sequential attitude priming paradigm*—tests the hot cognition hypothesis. Figure 2 schematically depicts variants of the two priming paradigms.

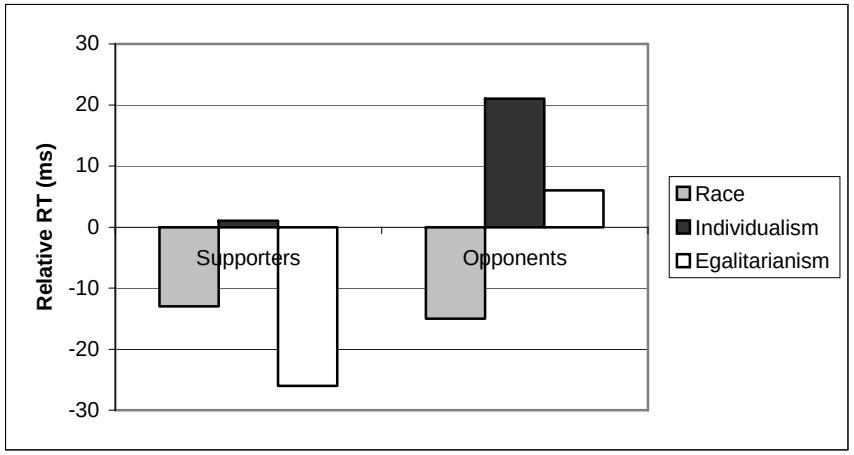
In the generic paradigm (Panel a) a prime word is presented followed by a target word. Critical here for demonstrating automaticity is to carefully control the time from the onset of the prime to the onset of the target so that the prime appears too quickly for conscious awareness and the target appears at the prime's peak of activation. Exposure times for primes under 200 ms guarantee that subjects will not be able to consciously identify the prime word. To precisely control the exposure time of the prime it is "masked" both before and after with meaningless symbols that erase all images from the visual cortex.

In the classic word/nonword lexical decision task (Panel b), experimental subjects are asked to categorize "as fast as possible without making too many errors" whether the letters form a word (e.g., DIAMOND) by pressing a Yes button or pressing No for nonwords (e.g., MADDION). The speed with which they make this lexical judgment measures the momentary accessibility of that concept in memory. When as depicted in Panel b, we manipulate the associative pathways through priming, this procedure measures the cognitive associations that come spontaneously to mind in response to the prime word. For example, presenting the prime JEWELRY—even at subliminal speeds—will reliably facilitate the Yes response to DIAMOND relative to a baseline control. The faster the prime-target response time the closer the cognitive association.

For example, we could use the lexical decision task to measure the preconscious associations that come automatically to mind when people think about racially-charged political issues. Do people think of ideological considerations, first and foremost, or do racial attitudes and stereotypes dominate? To put this debate between principled conservatism and symbolic or modern racism to the test (Panel c), we measured the strongest spontaneous associations of supporters and opponents to the concepts Affirmative Action and Welfare. Experimental participants categorized as words (button press Yes) or nonwords (No) three sets of pretested targets: racial stereotype words (e.g., afro, rap, basketball), individualism words (e.g., earn, merit, hand-outs), and egalitarian-

ism words (e.g., equality, opportunity, need). These target words were primed with AFFIRMATIVE ACTION, WELFARE, a random set of race-irrelevant primes, or by semantically meaningless baseline primes (e.g., PPPPPP), all presented outside of conscious awareness at the subliminal speed of 39 ms. The research question

Figure 3: Automatic Associations to Affirmative Action



is whether priming race policies will facilitate or inhibit the lexical categorization of ideological or racial stereotype words.[32]

Figure 3 reports the average response time (RT) for this lexical task on the three sets of target words when primed with Affirmative Action minus the average baseline RT for the same targets when primed by the nonsense Baseline (so negative bars represent facilitation and therefore an association in mind between Affirmative Action and the target). Contrary to the expectations of principled conservatism, opponents of affirmative action thought first and foremost about race and not at all about ideological principle; indeed, racial stereotypes were activated preconsciously for supporters and opponents alike when primed with Affirmative Action and Welfare, while principle (egalitarianism) was activated only for supporters.

[32] Burdein and Taber 2004. For the debate over the underpinnings of support or opposition for race policies like affirmative action, see Kinder and Sears 1981; Sears, Sidanius and Bobo 2000; Sniderman and Tetlock 1986.

But what about affect? Are one's feelings also activated when a concept is primed? That is the hot cognition question central to our *JQP* model of political information processing. To turn the hot cognition conjecture into a testable hypothesis, the *sequential attitude priming paradigm* was developed for empirically testing the postulate that feelings are directly linked to concepts in memory and are spontaneously activated on their mere exposure. As with the lexical decision task, this procedure exposes subjects to a masked prime followed by a target, and we are interested in the facilitation (speed up) or inhibition (slow down) effect of the valence of the prime on processing of the target. But here—as depicted in Figure 2, Panel d—the subject's task is to press a button labeled "positive" or "negative" to indicate "as fast as possible without making too many errors" whether the target word, chosen for its unambiguous positive or negative meaning (e.g., "sunshine" or "cancer"), has a positive or negative connotation. The hot cognition hypothesis predicts facilitation for *affectively congruent* prime-target pairs (COCKROACH would speed the evaluation of RABIES) and inhibition for *incongruent* pairs (COCKROACH would slow down the time to indicate RAINBOW is a positive word). Note that this is a non-reactive task: we never explicitly ask people to indicate whether the target describes a cockroach nor whether they feel positive or negative about cockroaches, but rather to simply indicate whether the target word is positive or negative.[33] While there are now hundreds of attitude priming studies in the psychological literature and strong experimental support for the spontaneous processing of affective information, there have been few tests of the automaticity of affect for political leaders, groups, and issues, and fewer still explorations of individual differences and situational constraints on the automaticity of attitudes in the political domain. To address these concerns we carried out three

[33] Bargh, Chaiken, Govender and Pratto 1992; Fazio, Sanbonmatsu, Powell and Kardes 1986.

priming experiments designed to (1) test for the generalizability of hot cognition to political leaders, groups, and issues, (2) test for the automaticity of affect by employing the subliminal version of the attitude priming paradigm to guarantee that processing is at the preconscious level, and (3) test whether political sophistication and attitude strength moderate the automaticity of political attitudes.[34] In Table 1 we list the primes and targets used in one of our studies.

Table 1: Primes and Targets for Hot Cognition Study

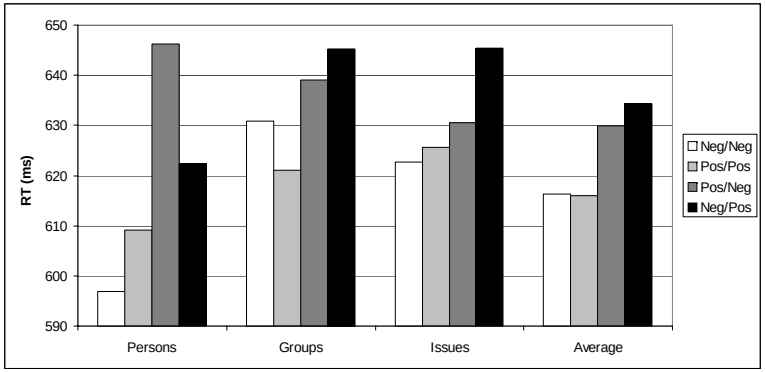
Person Primes	Group Primes	Issue Primes
Colin Powell	African Americans	Affirmative Action
George W. Bush	Americans	Counter-Terrorism
Giuliani	Arabs	Death Penalty
Hillary	Democrats	Free Speech
Hitler	Jews	Gun Control
Kennedy	NAACP	Pro-Choice
Lincoln	NRA	Pro-Life
Mark Green	Politicians	Taxes
Mike Bloomberg	Republicans	Welfare
Osama bin Laden	Terrorists	
Pataki		
Positive Targets	**Negative Targets**	
joy	death	
laughter	demon	
rainbow	grief	
gift	pain	
hug	rabies	

Figure 4 presents the results from our Study 3 in sets of 4 bars, each representing the mean response time (RT) for one of the four groups defined by the prime by target valence interaction:

[34] Lodge and Taber 2005.

from left to right, negative primes when primed with negative targets, positive primes/positive targets, positive primes/negative targets, and negative primes/positive targets. The hot cognition hypothesis predicts that the response times to attitudinally congruent prime-target pairs will be faster (facilitation) than to attitudinally incongruent pairs (inhibition). The informative comparison is between the first and third bars (for negative targets) and between the second and fourth bars (for positive targets).

Figure 4: Hot Cognition for Political Persons, Groups, and Issues

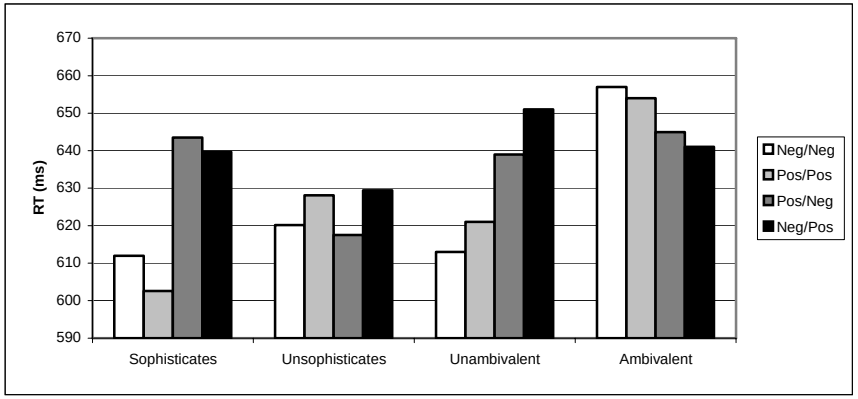


The key result is captured in Figure 4, which shows the expected pattern of facilitation and inhibition effects, broken down by prime type. Averaged responses across a wide range of political primes—current and historical political figures, political parties and other groups, as well as a variety of political issues—show clear evidence of an automatic link in memory between political concepts and positive or negative affect. These feelings are evoked spontaneously without the citizen having to consciously conjure up the specific considerations that are routinely modeled as causing the affective evaluations.

Our *JQP* model also predicts that these automatic facilitation and inhibition effects will be greater for political sophisticates, who have thought about and repeatedly evaluated political leaders, groups, and issues. Citizens with below average interest and politi-

cal knowledge are unlikely to have formed strong affective links in memory for many of these political primes (e.g., Mark Green), and therefore should not display the pattern of facilitation and inhibition that indicates hot cognition.

Figure 5: Sophistication and Ambivalence Effects in Hot Cognition



In Figure 5, we see the sophistication results for our Study 3 averaged across all target types (persons, groups, issues). As expected, sophisticates showed automatic affective responses to leaders, groups, and issues, while unsophisticates did not. This overall pattern lends credence to the theoretical expectations underlying the formation of online tallies and suggests that sophisticates, because of their interest in politics, have formed crystallized attitudes to a broader range of political objects than have unsophisticates, and these feelings come spontaneously to mind on mere (even subliminal) exposure to the concept.

Another boundary condition constraining the automaticity of attitudinal responses may be the ambivalence of some political attitudes, which may require a different processing mechanism than simpler univalent attitudes.[35] Recall in Figure 1 we represented ambivalent attitudes (tax cuts, e.g.) as having links to both positive and negative affect. In theory, priming an ambivalent attitude object would pass activation to both positivity and negativity, which should dampen any facilitation or inhibition effects on subsequent judgments. As shown in Figure 5, we do indeed find automatic inhibition and facilitation effects for unambivalent primes, but not for ambivalent primes, which strongly suggests that hot cognition effects are limited to those concepts which evoke strong and relatively univalent feelings.

[35] Lavine 2001.

In short, we found compelling evidence for the automaticity of affect across a broad range of political concepts, including political leaders, groups, and issues. As expected, these effects were strongest for sophisticated respondents and for unambivalent attitude objects. Feelings, it now appears, are directly linked to many concepts in memory, and can be spontaneously aroused, even when one is consciously unaware of the stimulus. These feelings are then available to guide subsequent processing and behavior. Indeed, they are unavoidable.

The Citizen as Motivated Reasoner

The deliberations of citizens are always motivated by some mix of the sometimes contrary desires to see the world accurately and as they would like it to be. Conventional models of political reasoning have emphasized the former, taking it largely on faith that citizens conscientiously think about the choices before them. Alas, we have already seen strong evidence that a variety of unnoticed, automatic forces intrude to disrupt the fable of the deliberatively rational citizen. The sophisticated citizen on the eve of an election is unlikely to be evenhanded when reading the papers, watching TV reports, or browsing the internet, for she already knows and *feels* her liking for the candidates, parties, and issues. Moreover, as a result of selective processing of the evidence, the more she thinks and deliberates, the stronger she will feel and the more extreme will be her attitudes.

In our theory, judgmental biases and attitude polarization are triggered by an initial (and uncontrolled) affective response. Motivated reasoning is a result of *hot* cognition. As soon as a citizen

recognizes the "subject, object, and verb" of a news report—that is, who is saying or doing what to whom—the feelings associated with these objects come inescapably to mind, whether consciously recognized or not, and for better or worse these feelings guide subsequent thought. Though not impossible, it is difficult to overcome the momentum of one's priors, especially when they are strong and based on an online integration of the considerable information a sophisticate will have processed over the course of the campaign. Is this rational? Sometimes. On the one hand, the online distillation of the history of a campaign into the affect one feels toward the candidates provides a remarkably efficient answer to the paradox of the minimal citizen. On the other hand, when citizens habitually avoid uncomfortable truths, and mold the evidence to suit their priors, it is hard to see how the marketplace can become a meeting place of ideas.

But what evidence is there for motivated biases in political information processing? We conducted a series of experiments to test the central motivated reasoning hypotheses in the context of political deliberation.[36] We were particularly interested in testing for *selective exposure*, where people seek out supportive arguments, a *confirmation* bias, where people treat evidence that supports their priors as stronger, and a *disconfirmation* bias where people accept supportive evidence uncritically but actively counterargue challenging evidence, the net effect of which will be attitude *polarization*. We expect to find these biases among the more sophisticated citizens and those with the strongest prior attitude, whereas unsophisticates and those with weaker attitudes will lack the motivation and/or knowledge necessary to defend their attitudes.

To test selective exposure, we ran and then replicated the following simple experiment: first we measured participants' prior feelings on two issues, gun control and affirmative action, then gave participants the opportunity to select pro or con arguments on these issues, using a computerized information board where subjects could click a labeled box to open an argument (e.g., arguments against gun control were labeled as coming from the NRA and arguments for affirmative action were labeled NAACP), and after they opened and read eight arguments (with some other intervening material), we again measured their attitudes on these two issues. As predicted by the selective exposure hypothesis, participants—especially political sophisticates—were significantly more likely to read the arguments of sympathetic sources than to expose themselves to an opposing point of view. Moreover, they polarized as a result of their selective exposure: subjects who were most biased in their information search became more extreme in their attitudes, while subjects below the median in search bias did not polarize.

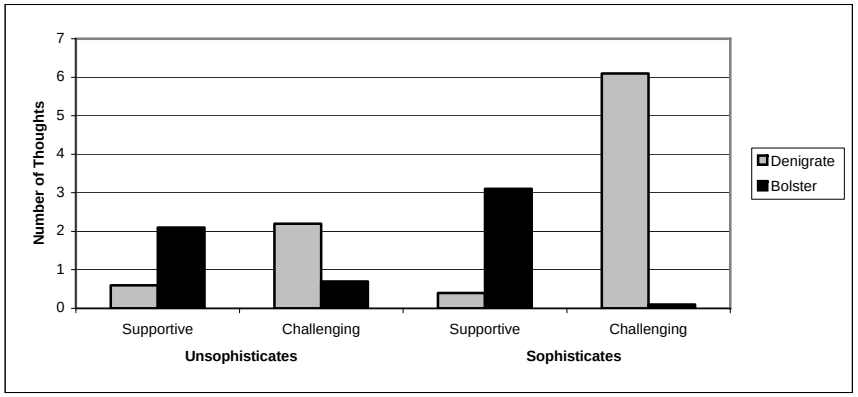
[36] Lodge and Taber 2000; Taber and Lodge 2006; Taber, Lodge and Glather 2001.

We also included several tests of confirmation and disconfirmation bias in our experiments. First, we presented to subjects in random order a set of four pro and four con arguments on either affirmative action or gun control, and asked them to rate the strength of each. Our instructions were designed to promote evenhandedness, but we did not expect that sophisticates and those with strong priors would treat both sides equally. They did not: in both studies and across two political issues, sophisticates and those with strong priors showed significant bias in rating the arguments with which they agreed as stronger than those with which they disagreed. Moreover, our subjects spent considerably more time rating the counter- than the pro-attitudinal arguments, which suggests that they may have been actively counterarguing. To test this possibility more directly, we asked our subjects to list the thoughts that came to mind as they rated half of the arguments (2 pro, 2 con). The disconfirmation hypothesis was again supported, as Figure 6 demonstrates. For all subjects, but especially for those high in political knowledge, the vast majority of thoughts were attitude consistent: either attempts to bolster supporting arguments or to denigrate and counterargue challenging arguments.

Like selective exposure, these processing biases polarized attitudes. Subjects whose ratings of the 8 arguments were above the sample median on our measure of disconfirmation bias (average

rating of consistent arguments minus average for inconsistent arguments) polarized nearly 24% for affirmative action and over 16% for gun control, while those below the median did not polarize.

Figure 6: Disconfirmation Biases in Rating Political Arguments



These studies show that people find it very difficult to escape the pull of their spontaneously evoked feelings. First, people simply feel that the information they agree with is stronger than the information with which they disagree. Second, when thinking about the evidence on a policy issue, people actively denigrate the information with which they disagree while accepting supportive information with little scrutiny. Third, people seek out confirmatory information and avoid evidence that might challenge their priors. Fourth, all of these biases conspire to drive attitudes further in the direction of priors the more they think and reason about the issues. Finally, all of these biases are particularly pronounced for citizens with more knowledge and stronger political attitudes, the very folks on whom democratic theory relies most heavily.

Our *John Q. Public* model posits that motivated reasoning—the systematic biasing of judgments in favor of one's immediately accessible beliefs and feelings—is built into the basic architecture

and information processing mechanisms of the brain.[37] Because the spreading of activation for both cognitive and affective associations operates below conscious awareness, the deliberative construction of beliefs, attitudes, and intentions necessarily involves the integration of the conscious and unconscious associations that are momentarily accessible. Hence, when called on to recount or justify an opinion, there is likely to be a disjuncture between implicit and explicit measures of beliefs and attitudes, for four reasons:

1. Whereas long term memory is vast, the sample of considerations that enter conscious working memory is severely limited to 7 ± 2 concepts and their affective tags. Given the paucity of cognitive associations available for conscious consideration (e.g., the modal number of likes and dislikes reported by NES respondents for their liking of congressional candidates is one), a respondent asked why she favors or opposes a candidate or issue has three options: she can say "I don't know why I feel this way," she can report one or two associations that come easily to mind, or—if pressed for an answer—she can rationalize, that is, dredge up some reasonable explanation for why one would or should feel this way.
2. By contrast with the severe limitations on conscious thought, a staggering amount of information can be processed unconsciously. We can process about 11 million bits of information per second on a nonconscious level; when we attempt conscious mental calculations (e.g., an arithmetic problem), we can handle about 12 bits per second. Most thinking is unconscious, by at least six orders of magnitude![38]
3. The affective tags linked to social concepts in memory decay at a slower rate than do the considerations that originally entered into the evaluation. For example, Lodge, Steenbergen, and Brau found that the facts and considerations that originally informed the evaluation of a political candidate decayed

[37] Gazzaniga 1992; Gazzaniga 2000.
[38] Hassin, Uleman and Bargh 2005.

at an exponential rate, while the affective evaluation of the candidate remained stable over 31 days.[39]

4. The sampling of considerations from LTM is not likely to be a random draw, but will be biased in favor of affectively congruent associations.
5. As shown in all pre and postconscious automaticity experiments, with the "backpack" versus "briefcase" study a case in point, judgments are very sensitive to contextual factors, and consequently, as natural storytellers people will generate more plausible than veridical accounts for how and why they say and act as they do.[40]

While the general principles guiding the role of accessibility and retrieval of information are well-known, the implicit versus explicit distinction goes to the heart of our discipline's explanations for how, when, and why citizens think, reason, and act as they do. We expect that people will routinely rely on their spontaneous thoughts and feelings unless confronted by irrefutable evidence, social pressure, or challenges to self-images. The experimental literature presents clear evidence that automatic processes underlie *all* conscious processing and are especially powerful determinants of evaluations when: hot cognitions are available and strong; explicit measures are tainted by social desirability, deceit, or prejudice; the costs of being wrong are low; attentional resources are otherwise engaged or distracted; one is under time pressure; an environmental event is noticed but not recognized as being influential; and one's behavior is not so consequential as to trigger questions about "why did I think, feel, say, or do that?"

These situational and contextual factors favoring automaticity appear to characterize the world of politics for most citizens most of the time, where, typically, the direct consequences of political beliefs and attitudes are distant and indirect, where uncertainty reigns, rumination is rarely called for, where one is easily distracted by rapid-fire TV images, and where the stream of information

[39] Lodge, Steenbergen and Brau 1995.

[40] Clore and Isbell 2001.

parallels one's thoughts with congenial cues. It takes concerted effort to change habits of thinking, feeling, and doing, especially when these automatically-evoked processes are at least partly hard-wired.

At this juncture we have documented incidental priming effects on judgments, evaluations, and behavior and we have shown effects of automatic feelings on the processing of political arguments. Virtually all the studies reported here, and many more from the experimental literature, document the immediate consequences of some affective event—whether conscious or not—on evaluations or behavior. Sometimes there is a feeling of unease with the considerations that come to mind, or a sensed dissociation of implicit from explicit thoughts, feelings, and actions. If consciously conflicted, one may make the effort to resolve the dissonance among and between thoughts and feelings. But there is now reason to believe that spontaneously activated priming effects are difficult to correct for, and especially so when people stop, think, and actively engage in deliberative processing (Forgas, 1995; Wilson, 2002). First, because the search of memory favors chronically associated pathways, both affective and cognitive, one is typically not aware of a conflict among or between beliefs and attitudes. Being congruent, the thoughts, feelings, and justifications that come spontaneously to mind often "feel" right, and do not produce a sense of dissonance. Second, because we are only consciously aware of the outputs of our thinking and not the processes that link thoughts to feelings to evaluations, we are prone to fill in the blanks with plausible reasons for the associations that came to mind. Because of the prior attitude bias in memory search, the greater the conscious effort to query memory when constructing a response, the likelier the sample of retrieved considerations will be skewed in favor of affectively congruent considerations.[41]

In short, we expect that unconsciously processed priming effects bias downstream thinking and reasoning by activating congruent pathways in memory which skews the sample of considera-

[41] Forgas 1995; Forgas 2000.

tions that come to mind and promotes attitude polarization. One recent study, for example, found evidence that simple affective primes (smiley or frowney cartoon faces, presented at subliminal speeds) influenced the thoughts that came to mind about illegal immigration, which in turn significantly influenced subjects' evaluations of contemporary policy proposals to deal with illegal immigration.[42]

Conclusion

We see two fundamental implications of this research on unconscious thought processes for the practice of contemporary political science, both of which will require that we dramatically rethink our notions of how people mentally represent the world and how we go about measuring and modeling citizens' expressions of political beliefs and attitudes. The first, documented throughout this essay, is how hot cognition kick-starts conscious and unconscious processes that promote response biases, especially when citizens encounter incongruent information about political candidates, groups, and issues. While these experimental findings do not suggest that people, even the much-maligned American voter, are compelled to act contrary to their explicit beliefs and conscious choices, the evidence is strong that the monitoring and correction of unconscious biases is difficult at best.[43] Moreover, because deliberation is necessarily influenced by prior, unconscious factors to a degree and in ways that are inaccessible to introspection, we must be highly suspicious of self-reported beliefs, attitudes, and behavioral intentions.

In our view, one of the more interesting—and paradoxical—implications of *John Q. Public* is that people internalize simple summary evaluations, formed spontaneously online as they encounter political information. Once formed, such running tallies provide decision makers a ready-made, highly accessible likeability

[42] Erisen, Lodge and Taber 2007.

[43] Devine 1989; Ditto and Lopez 1992; Taber and Lodge 2006.

heuristic to inform their beliefs and guide their behavior.[44] Citizens need not rely on fallible memory traces of the considerations that originally informed their attitudes in order to act on their preferences. Thus we find that citizens feel their preferences far faster and more accurately than they can explicitly report the reasons for their beliefs and attitudes. A central postulate of our model is that online tallies are sensitive to and keep account of the costs and benefits of one's experiences, and their automatic activation is among the most powerful heuristics informing thought and action. From this vantage point, top-of-the-head judgments and evaluations may prove to be more reliable guides to the expression of attitudes and behavior than those based on the careful consideration of pros and cons.

A second obvious implication is for rational choice theories of political action. Our field has not been kind to the democratic citizen. Normative democratic theory imposes heroic expectations about the capacity and motivation of *homo politicus*, while modern empirical research finds many citizens to be *not-so-sapien*. Surveys consistently find respondents to be distressingly ignorant of things political, suggesting at a minimum that many citizens are unable to form coherent attitudes, impressions, and evaluations, or choose intelligently among political leaders, groups, and ideas.

Interestingly, *JQP*'s automaticity-of-affect heuristic shares a common assumption with rational choice theory. To wit: rational decision makers aim to take actions that promote positive and avoid negative outcomes. This the case, anticipated feelings activated by the choices before us drive rational behavior—decision makers choose that option they feel will maximize the net balance of positive over negative emotions, essentially saying "I'll be happier if I choose candidate A over B". From this *consequentialist* perspective, utility maximization is defined as taking actions expected to promote positive emotions. The elicitation of positive or negative feelings is at the core of decision making, whether rational, satisficing, or thoughtless. It is the anticipation of a more positive outcome (by our account experienced as a positive feeling via a

[44] Sears 2000; Sniderman, Brody and Tetlock 1991; Taber 2003.

somatic marker) that drives both psychologically realistic models of decision making and rational choice models.[45]

From this perspective, a key to rational behavior is the decision maker's ability to first access and then forecast the strength and direction of feelings toward the choices at hand. Leaving aside the well-documented inability of people to appreciate the probability of outcomes, we part company with the EU model on three counts: First, because feelings of positivity or negativity are activated before conscious considerations come into play, the appraisal of options is influenced preconsciously by the anchoring of prior feelings, which facilitates a biased search of memory, which then promotes rationalization processes that systematically fill in the gaps between feelings and thoughts. And second, affect enters into the decision stream at every step of the process, not only during the appraisal of choices. Rational models of judgment and choice fall short because of their general failure to model the effects of spontaneous feelings on the encoding, interpretation, search and retrieval of outcomes, evaluation of the desirability of alternatives, and the processes involved in anticipating the emotional consequences of choices. In addition to the emotions stemming directly from the appraisal of expected outcomes, these indirect effects impact the immediate decision process. Random and systematic biases fill the void left by the loss of episodic knowledge: random errors occur because contextual details are quickly forgotten, systematic biases because belief-consistent information is more likely to be recalled, and when thoughts and feeling are congruent it is time to end the search.

Finally, unless the appraisal of the anticipated consequences of one's options is accompanied by somatic feedback signaling positive or negative reaction, there is no "body loop" feeling linked to the options. Absent this somatic marker, people cannot effectively forecast their emotional responses to anticipated outcomes. While all social concepts and cognitive processes are affectively charged, there is a gap between the conscious appraisal of considered options (essentially a semantic response) and the emotional

[45] Loewenstein, Weber, Hsee and Welch 2001.

experience. People in one emotional state are not good at predicting their own or others' emotional reactions in a different state. This disjuncture between "is" and "as-if" experiences is especially problematic when the considered choices are calculated in a cool, deliberative way, as when survey respondents are called on to report their past, present, or future responses.

At this juncture then we are skeptical of the ability of citizens to reliably and veridically access the sources of their beliefs, the reasons for their attitudes, and their past or future intentions and actions. Much if not most of our experience takes place outside our conscious awareness, and as our recollections fade from memory they are replaced by socially constructed rationalizations about how and why we as well as others are likely to behave. Which automatic responses are activated depends on the set of preconditions operative in the environment *at the moment* and what's going on inside the individual's head *at the moment*. The key here is that once triggered, once the preconditions come into play, thoughts are linked to feelings to goals to choices outside of conscious awareness without necessarily triggering conscious or deliberative guidance. Where, when, how, and for whom deliberative processing will successfully override the automatic response is the critical, heretofore unanswered question that goes to the heart of all discussions of human rationality and the meaning of a responsible electorate.

References

Anderson, Norman H., and Alfred A. Barrios. 1961. Primacy Effects in Person Impression Formation. *Journal of Abnormal Social Psychology* 43 346-50.

Asch, Solomon E. 1946. Forming impressions of personality. *Journal of Abnormal and Social Psychology* 41 (3): 258-90.

Bargh, John A. 1997. The automaticity of everyday life. In *Advances In Social Cognition*, Ed. Robert Wyer, 1-61. Mahwah, NJ: Erlbaum.

Bargh, John A., Shelly Chaiken, Rajen Govender, and Felicia Pratto. 1992. The Generality of the Automatic Attitude Activation Effect. *Journal of Personality and Social Psychology* 62 (6): 893-912.

Bargh, John A., Mark Chen, and Lara Burrows. 1996. Automaticity of Social Behavior: Direct Effects of Trait Construct and Stereotype Activation On Action. *Journal of Personality and Social Psychology* 71 (2): 230-44.

Bechara, A., H. Damasio, D. Tranel, and A. R. Damasio. 1997. Deciding advantageously before knowing the advantageous strategy. *Science* 275 (5304): 1293-95.

Betsch, Tilman, Henning Plessner, Christine Schwieren, and Robert Gutig. 2001. I Like It But I Don't Know Why: A Value-Account Approach to Implicit Attitude Formation. *Personality and Social Psychology Bulletin* 27 (2): 242-53.

Burdein, Inna, Milton Lodge, and Charles S. Taber. 2006. Experiments On the Automaticity of Political Beliefs and Attitudes. *Political Psychology* 27 (3): 359-71.

Burdein, Inna, and Charles S. Taber. 2004. What Are They Thinking? Implicit Measures of Ideology and Racism. In *Annual Meeting of the Midwest Political Science Association*. Chicago, IL.

Chen, M., and J. A. Bargh. 1999. Consequences of automatic evaluation: Immediate behavioral predispositions to approach or avoid the stimulus. *Personality and Social Psychology Bulletin* 25 (2): 215-24.

Clore, Gerald , and Linda Isbell. 2001. Emotions as Virtue and Vice. In *Political Psychology in Practice*, Ed. James H. Kuklinski, 103-23. Hillsdale, CA: Lawrence Erlbaum Associates.

Crites, Stephen L., John T. Cacioppo, Wendi L. Gardner, and Gary G. Berntson. 1995. Bioelectrical Echoes from Evaluative Categorization: A Late Positive Brain Potential That Varies as a Function of Attitude Registration Rather Than Attitude Report. *Journal of Personality and Social Psychology* 68 (6): 997-1013.

Damasio, Antonio R. 1994a. *Descartes' Error: Emotion, Reason, and the Human Brain*. New York: G.P. Putnam.

Damasio, Antonio R. 1994b. Descartes Error and the Future of Human Life. *Scientific American* 271 (4): 144-44.

Damasio, Antonio R. 1999. *The Feeling of What Happens: Body and Emotion in the Making of Consciousness*. New York ; San Diego: Harcourt Brace.

Devine, Patricia. 1989. Stereotypes and Prejudice: Their Automatic and Controlled Components. *Journal of Personality and Social Psychology* 56 680-90.

Dijksterhuis, A., and H. Aarts. 2003. On wildebeests and humans: The preferential detection of negative stimuli. *Psychological Science* 14 (1): 14-18.

Dijksterhuis, Ap, and Ad van Knippenberg. 1998. The Relation Between Perception and Behavior, Or How to Win a Game of Trivial Pursuit. *Journal of Personality and Social Psychology* 74 (4): 865-77.

Ditto, Peter H., and David F. Lopez. 1992. Motivated Skepticism: Use of Differential Decision Criteria for Preferred and Nonpreferred Conclusions. *Journal of Personality and Social Psychology* 63 568-84.

Erisen, Cengiz, Milton Lodge, and Charles S. Taber. 2007. The Role of Affect in Political Deliberation. In *The Annual Meeting of the American Political Science Association*. Chicago, IL.

Fazio, R. H., D. M. Sanbonmatsu, M. C. Powell, and F. R. Kardes. 1986. On the Automatic Activation of Attitudes. *Journal of Personality and Social Psychology* 50 (2): 229-38.

Forgas, John. 1995. Mood and Judgment: The Affect Infusion Model. *Psychological Bulletin* 117 (1): 39-66.

Forgas, John. 2000. Feeling and Thinking: The Influence of Affect on Social Cognition and Behaviour. *International Journal of Psychology* 35 (3-4): 109-09.

Gazzaniga, Michael S. 1992. *Nature's mind : the biological roots of thinking, emotions, sexuality, language, and intelligence*. New York: BasicBooks.

Gazzaniga, Michael S. 2000. *The new cognitive neurosciences*. Cambridge, Mass.: MIT Press.

Greenwald, Anthony G., and Mahzarin R. Banaji. 1995. Implicit Social Cognition - Attitudes, Self-Esteem, and Stereotypes. *Psychological Review* 102 (1): 4-27.

Hassin, Ran R., James S. Uleman, and John A. Bargh. 2005. *The New Unconscious*. Oxford ; New York: Oxford University Press.

Hastie, Reid, and Bernadette Park. 1986. The Relationship between Memory and Judgment Depends on Whether the Judgment Task Is Memory-Based or Online. *Psychological Review* 93 (3): 258-68.

James, William. 1890. *The principles of psychology*. New York,: H. Holt and company.

Kay, Aaron C., S. Christian Wheeler, and John A. Bargh. 2004. Material Priming: The Influence of Mundane Physical Objects On Situational Construal and Competitive Behavioral Choice. *Organizational Behavior and Human Decision Processes* 95 (1): 83-96.

Kim, Sung-youn. 2005. A Model of Political Information Processing: The Dynamics of Candidate Evaluation and Strategic Behaviors in Social Dilemma Games. In *Political Science*. Stony Brook, NY: Stony Brook University.

Kim, Sung-youn, Charles Taber, and Milton Lodge. 2007. A Model of Survey Response: Perceptions about Candidate Issue Positions. In *Annual Meeting of the American Political Science Association*. Chicago, IL.

Kinder, Donald R., and David O. Sears. 1981. Prejudice and Politics: Symbolic Racism Versus Racial Threats to the Good Life. *Journal of Personality and Social Psychology* 40 (3): 414-31.

Kunda, Ziva. 1990. The Case for Motivated Reasoning. *Psychological Bulletin* 108 (3): 480-98.

Lavine, Howard. 2001. The Electoral Consequences of Ambivalence Toward Presidential Candidates. *American Journal of Political Science* 45 (4): 915-29.

LeDoux, Joseph E. 1996. *The Emotional Brain: The Mysterious Underpinnings of Emotional Life*. New York: Simon & Schuster.

Lodge, M., K. M. Mcgraw, and P. Stroh. 1989. An Impression-Driven Model of Candidate Evaluation. *American Political Science Review* 83 (2): 399-419.

Lodge, Milton, Marco R. Steenbergen, and Sean Brau. 1995. The Responsive Voter - Campaign Information and the Dynamics of Candidate Evaluation. *American Political Science Review* 89 (2): 309-26.

Lodge, Milton, and Charles S. Taber. 2000. Three Steps Toward a Theory of Motivated Political Reasoning. In *Elements of Reason*, Ed. Arthur Lupia, Matthew McCubbins and Samuel Popkin, 183-213. Cambridge, UK: Cambridge University Press.

Lodge, Milton, and Charles S. Taber. 2005. The Automaticity of Affect for Political Leaders, Groups, and Issues: An Experimental Test of the Hot Cognition Hypothesis. *Political Psychology* 26 (3): 455-82.

Loewenstein, George. F., Elke U. Weber, Christopher K. Hsee, and Ned Welch. 2001. Risk as Feelings. *Psychological Bulletin* 127 (2): 267-86.

Marcus, George E., W. Russell Neuman, and Michael MacKuen. 2000. *Affective Intelligence and Political Judgment*. Chicago: University of Chicago Press.

Mendelberg, Tali. 2001. *The race card : campaign strategy, implicit messages, and the norm of equality*. Princeton, N.J.: Princeton University Press.

Morris, James P., Nancy K. Squires, Charles S. Taber, and Milton Lodge. 2003. Activation of Political Attitudes: A Psychophysiological Examination of the Hot Cognition Hypothesis. *Political Psychology* 24 (4): 727-45.

Mullen, Brian, David Futrell, Debbie Stairs, Dianne Tice, Roy Baumeister, Kathleen Dawson, Cataherine Riordan, Christine Radloff, George Goethals, John Kennedy, and Paul Rosenfeld. 1986. Newscaster's facial expressions and voting behavior. Can a smile elect a president? *Journal of Personality and Social Psychology* 51 291-95.

Nosek, Brian, and Frederick Smyth. 2007. A Multitrait-Multimethod Validation of the Implicit Association Test: Implicit and Explicit Attitudes are Related but Distinct Constructs. *Journal of Experimental Psychology* 54 (1): 14-29.

Schwarz, Norbert, and Gerald L. Clore. 1983. Mood, misattribution, and judgments of well-being: Informative and directive functions of affective states. *Journal of Personality and Social Psychology* 45 513-23.

Sears, David O. 2000. The Role of Affect in Symbolic Politics. In *Citizens and Politics: Perspectives from Political Psychology*, Ed. James H. Kuklinski, 14-40. Cambridge, UK: Cambridge University Press.

Sears, David O., Jim Sidanius, and Lawrence Bobo. 2000. *Racialized Politics: The Debate About Racism in America*. Chicago: University of Chicago Press.

Sigelman, Lee. 1990. Toward a stupidity-ugliness theory of democratic elections and debacles. *PS: Political Science and Politics* 23 (1): 18-20.

Sniderman, Paul, Richard Brody, and Philip Tetlock. 1991. *Reasoning and Choice: Explorations in Political Psychology*. Cambridge, UK: Cambridge University Press.

Sniderman, Paul, and Philip Tetlock. 1986. Symbolic Racism: Problems of Motive Attribution in Political Analysis. *Journal of Social Issues* 42 (2): 129-50.

Taber, Charles S. 2003. 2003. Information Processing and Public Opinion. In *Handbook of Political Psychology*, Ed. David O. Sears, Leonie Huddy and Robert L. Jervis, 433-76. London: Oxford University Press.

Taber, Charles S., and Milton Lodge. 2006. Motivated Skepticism in the Evaluation of Political Beliefs. *American Journal of Political Science* 50 (3): 755-69.

Taber, Charles S., Milton Lodge, and Jill . 2001 Glather. 2001. The Motivated Construction of Political Judgments. In *Citizens and Politics: Perspectives from Political Psychology*, Ed. James H. Kuklinski, 198-226. London: Cambridge University Press.

Thagard, Paul. 2006. *Hot Thought: Mechanisms and Applications of Emotional Cognition*. Cambridge, MA: MIT Press.

Uleman, James S., and John A. Bargh. 1989. *Unintended thought*. New York: Guilford Press.

Weinberger, Joel, and Drew Westen. 2007. RATS, we should have used Clinton: Subliminal priming in political campaigns. In *International Society of Political Psychology*. Portland, OR.

Wilson, Timothy D. 2002. *Strangers to ourselves : discovering the adaptive unconscious*. Cambridge, MA: Belknap Press of Harvard University Press.

Zajonc, Robert. 1980. Feeling and Thinking: Preferences Need No Inferences. *American Psychologist* 35 117-23.

Zaller, John, and Stanley Feldman. 1992. A Simple Theory of the Survey Response. *American Journal of Political Science* 36 579-616.