1
The Four Horsemen of Automaticity:
Awareness, Intention, Efficiency,
and Control in Social Cognition

John A. Bargh
New York University

Contents

The Decomposition of Automaticity 3
  Conditional Automaticity 4
  The Ecology of Automaticity 6
Awareness 7
  Subliminal Perception 8
  Misattribution 13
  Conclusions 15
Intentionality 16
  Automatic Attention Responses and Perceptual Selection 16
  Automatic Evaluation 17
  Automatic Stereotype Activation 21
  Spontaneous Trait Inference 23
Efficiency 24
  Social Judgment 25
  Dispositional Inference 26
  Conclusions 27
Controllability 28
  Situationally Induced Motivations 29
  Inernally Generated Motivations 29
  Conclusions 30
The Automaticity of Everyday Life: An Agenda for the Next 10 Years 30
Acknowledgments 31
References 31

I do not think, therefore I am.
—Jean Cocteau

When the first edition of this Handbook appeared in 1984, research on automatic phenomena was just beginning. In the 10 years preceding it, a total of 28 research articles were published on topics directly relevant to the automaticity of a social
psychological phenomenon. In the following 10-year-period there have been 123 such research articles. Clearly, that research on automatic phenomena in social psychology has mushroomed in the past decade.

There is now hardly a research domain or topic that has not been analyzed in terms of its automatic features or components. Much attention has been devoted to questions of whether dispositional inferences are made automatically, whether attitudes become activated automatically to influence ongoing behavior, whether accessible social constructs and stereotypes automatically affect one's judgments of oneself and others, whether people have automatic evaluative and emotional reactions to stimuli, and the degree to which a person is aware or unaware of the influences on his or her judgments and subjective experience.

In deciding how to structure a review of automaticity research, I faced a dilemma: Should it be organized in terms of specific content areas, such as attribution or stereotyping, and describe the extent to which these phenomena are found to be automatic in nature? This would be useful, except that it would miss many of the reasons why so much research attention has been given to questions of automaticity across different research domains. Those reasons have to do with the fact that the separate defining qualities of automaticity are important issues in their own right—the extent to which thought and behavior are unintentional, occur outside of awareness, are uncontrollable, and are efficient in their use of attentional resources.

Ten years ago, the consensus view (Johnson & Hasher, 1987) was that a mental process was either automatic—possessing all four of those qualities—or controlled, possessing all the opposite qualities (i.e., intentional, controllable, consumptive of limited attentional resources, and in awareness; see Bargh, 1984; Posner & Snyder, 1975; Shiffrin & Schneider, 1977). If a given process was not a member of one type, then by default it had to be a member of the other. Guided by this prevalent dichotomy, I argued at the time that many claims of automaticity within social psychology were not authentic, because they did not satisfy all four criteria.

---

1All of these are included in the Reference section. Because the earlier studies were reviewed in the first edition of this Handbook (Bargh, 1984), the present chapter focuses mainly on the post-1983 research. However, mention should be made here of those pioneering studies, and the following is, to my knowledge, a complete list of the pre-1983 research and theory directly relevant to one or more aspects of automaticity (and it is certainly possible that I missed some relevant articles): "top-of-the-head" attributions based on visual salience (Taylor, Crocker, Fiske, Spronk, & Winkler, 1979; Taylor & Fiske, 1978), that such salience effects occur automatically at encoding (Smith & Miller, 1979), behavior in routine social interactions (Langer, 1978; Langer, Blank, & Chanowitz, 1978), passive trait category priming effects on social judgment (Higgins, Rhose, & Jones, 1977; Snell & Wyer, 1979, 1980), the application of the self-representation in perceptual selection and encoding (Bargh, 1982; Getler & Shaver, 1976; Hull & Levy, 1979; Markus & Smith, 1981), one's lack of awareness of important influences on one's impressions and judgments (Lewicki, 1982; Nisbett & Wilson, 1977; Wegner & Vallacher, 1977), and one's frequent lack of awareness of the influential stimuli themselves (Bargh, 1982; Bargh & Pietromonaco, 1982; W. Wilson, 1979).
the possibility of motivational control (see Fiske, 1989), as well as a consideration of the separate stages of the stereotyping process and their differential controllability (Devine, 1989), showed that a process could be simultaneously unintended and efficient on the one hand, but nonetheless controllable. Therefore, the first moral of the present story is for researchers to be more specific about the particular qualities of automaticity they are demonstrating and claiming for the process in question—unintentionality, unawareness, uncontrollability, or high efficiency—instead of discussing only its automaticity or relative automaticity.

Conditional Automaticity

The second and related moral is that the various demonstrations of automatic processing in social cognition vary as to the conditions that are necessary for the process to occur. Some of the automatic phenomena that were identified required the person's intention for their initiation, others required substantial attentional support, others awareness of the triggering stimulus, and so on. In a previous analysis of social cognitive phenomena in terms of these conditions (Barth, 1989), three general sorts of automaticity could be identified: preconscious, postconscious, and goal-dependent.

Preconscious Automaticity. A preconsciously automatic process requires only that the person notice the presence of the triggering stimulus in the environment. These processes occur automatically when a stimulus is noticed, as part of the act of figural synthesis (Neisser, 1967), and do not require a deliberate goal or intention. Such processes include interpretations, evaluations, and categorization that occur prior to and in the absence of conscious or deliberative response to the stimulus (i.e., during the microgenesis of its perception; Werner, 1956). One certainly may be aware (and usually is) of the end result of this fast preconscious construction of the percept. Thus, preconscious is not synonymous with subliminal, although subliminal processes are certainly a subset of preconscious ones.

Examples of preconscious automaticity include chronically accessible trait constructs influence on social perception, because they occur without intention and even uncontrollably (Bargh & Pratto, 1986), as well as efficiently (Bargh & Thein, 1985). Automatic attitude activation also appears to qualify as a preconscious phenomenon, because it occurs without intention or controllability (Roskos-Ewoldsen & Fazio, 1992) and immediately and efficiently (Bargh, Chaiken, Govender, & Pratto, 1992; Fazio, Sanbonmatsu, Powell, & Kardes, 1986). Attitudes that are strong enough to become active automatically have been discussed in terms of the "chronic accessibility" of their association to the corresponding object representation (Fazio et al., 1986); therefore it is not surprising that both chronically accessible trait constructs and attitudes appear to share many preconscious automatic properties (see Bargh, 1984). Other forms of preconscious automaticity that have been documented are automatic attention responses to negative stimuli such as trait adjectives (Pratto & John, 1991) and angry faces (Hansen & Hansen, 1988), and physiological reactions to stimuli that are relevant to chronic concerns about the self (Strauman & Higgins, 1987). (It should be noted that many other phenomena ultimately may be found to be as unconditionally automatic as these, but the currently available experimental demonstrations of those phenomena include conditions, such as explicit instructions for subjects to engage in the process, that at the present time preclude conclusions about their unintentional nature [see Bargh, 1992b].)

Postconscous Automaticity. These effects are functionally the same as preconscious effects, except that they require some kind of recent conscious, attentional processing to occur. Priming effects on impression formation (e.g., Higgins, Rholes, & Jones, 1977; Srull & Wyer, 1979) are the best example. Other examples are repeated expression manipulations of attitude accessibility (Herr, Sherman, & Fazio, 1984) and the effect of a recent positive or negative experience (even as mild as having cookies in the subject waiting room) on the accessibility of positive versus negative life experiences (Isen, Shalker, Clark, & Karp, 1978).

Postconscious effects are functionally the same as preconscious ones, except that they are temporary instead of chronic and they result from the residual activation of conscious processing. For example, Fazio and his colleagues obtained the same results of accessible attitudes on behavior and attention whether the attitude was chronically (i.e., a preconscious effect) or temporarily accessible (e.g., Fazio et al., 1986; Roskos-Ewoldsen & Fazio, 1992), and studies comparing chronic and temporary construct accessibility show the same quality of effect for each (Bargh, Bond, Lombardi, & Tota, 1986; Bargh, Lombardi, & Higgins, 1988). The chronic versus temporary distinction between preconscious and postconscious processing is not a trivial one, however. Postconscious effects only occur given relevant thought and go away after a short time, preconscious effects are "eternally vigilant" (see Bargh, 1989; Bargh et al., 1988).

Methodologically, the phenomenon of postconscious automaticity—that temporary accessibility can mimic chronic accessibility—is a potential pitfall for researchers who intend to study unconditionally and chronic automatic effects. There have been several recent demonstrations of the effect of having subjects complete questionnaires prior to tests of how they think naturally or "automatically" in the same content domain. Skelton and Strohmetz (1990) showed that having subjects first consider common words for their health connotations resulted in a greater number of symptoms reported on symptom check lists. Mark, Sinclair, and Wells (1991) showed that giving subjects the Beck Depression Inventory (BDI) at the beginning of the experimental session produced different self-judgments by depressed versus nondisordered subjects compared to the condition in which the BDI was not administered first. Spielman and Bargh (1991) replicated two different studies that had reported automatic thought patterns in
depression, but had given subjects the BDI prior to the test of automaticity. In both studies, the original results were replicated only when subjects completed the BDI first.

These findings indicate that one must be careful not to prime or create postconsciously automatic phenomena by having subjects recently engage in a task that causes them to think about the same topics on which one is assessing their chronic or preconscious thought processes (Barth, 1990). Although similar effects are obtained in studies using priming or some other technique (e.g., repeated attitude expression) to create temporary accessibility as in studies of chronic accessibility, one cannot conclude that chronic, preconscious automaticity effects exist on the basis of demonstrations of temporary accessibility in that domain. Any mental representation or mode of thinking that is available in memory for use by the subject can be made accessible in an experiment, but this does not mean that every available mental structure or process is chronically accessible (see Higgins & King, 1981; Tulving & Pearlstone, 1966) and operates preconsciously.

Goal-Dependent Automaticity. The third general class of automatic phenomena only occurs with the person's consent and intent. Examples include the development of efficient behavior-to-trait judgments through practice (Smith & Bratscome, 1988; Smith & Lerner, 1986) and the evidence that self-concepts or other-concepts become active automatically given the intention to consider the self or another person (Barth & Tota, 1988; Dovidio, Evans, & Tyler, 1986; Perdue, Dovidio, Gurtman, & Tyler, 1990). For example, in the Barth and Tota study, negative trait concepts became active automatically (i.e., efficiently and immediately when under attentional load) when depressed subjects were asked to describe the self, but positive concepts were activated automatically when these subjects were trying to think about the average person. The same set of positive trait concepts were activated automatically in nondepressed subjects in both judgment contexts (see Pauly, Graf, & Van Selst, 1989; and Pauly & Levitt, 1987; for additional evidence of the increasing positivity of the self-concept in [nondepressed] subjects with increasing attentional load).

The Ecology of Automaticity

Decomposing the concept of automaticity into its component features in this way will also assist one to assess the ecological validity of the phenomenon in question. For example, suppose the effect requires that subjects be instructed to engage in such processing, as when they are given an explicit goal to form an impression or attribution. What is the likelihood that these subjects would spontaneously have that goal in their natural environment, in the absence of these situational demands? If an effect requires recent conscious thought relevant to the topic in question, how often will subjects normally be thinking along those lines? Clearly, to the extent that an effect does not require such preconditions, it will have a more frequent and important influence on thought, judgment, and behavior (Barth, 1992b).

One tradition of research in social psychology most closely identified with the work of Zajonc pursues the unconditional, "mere" effects of stimuli in this way. The mere presence theory of social facilitation (Zajonc, 1965), the mere exposure effect on attitude formation (e.g., Zajonc, 1968), the intellectual environment model of birth order effects on intellectual development (Zajonc & Markus, 1975), and the precognitive affective processing system (Zajonc, 1980) were all hypothesized to be unconditional mental phenomena. The hypothesis here—to push laboratory phenomena to their limits in exploring the minimum conditions necessary for their occurrence—is in the same tradition (see also McGuire, 1983).

Questions of the awareness, intentionality, controllability, and efficiency of thought and behavior are important in their own right, and transcend specific research domains. The issue of how much one is in control of one's thoughts and behavior was considered by Posner and Snyder (1975) to be a fundamental question of existence. Fiske (1989) pointed out the importance for the legal system of understanding the role of intentionality, because it strikes at the heart of the issues of responsibility and culpability for one's actions (e.g., in hiring and promotion discrimination cases). Gilbert (1991) argued that differences in how efficiently people accept versus question the validity of what they see and are told matter greatly in determining what they believe and their ability to guard against erroneous beliefs (see also Chandonit & Langer, 1981). In my opinion, these separate and distinct qualities of automaticity are important matters for study in and of themselves, not only as they are applied to specific research topics. In the following review of the literature on automaticity, I discuss the research in terms of its relevance for these issues of awareness, intentionality, efficiency, and control.

AWARENESS

There are three ways in which a person may be unaware of a mental process:

1. A person may be unaware of the stimulus itself, as in subliminal perception.
2. A person may be unaware of the way in which that stimulus event is interpreted or categorized, as stereotyping and construct accessibility research have demonstrated.
3. The person may be unaware of the determining influences on his or her judgments or subjective feeling states (e.g., the use of felt ease of perceptual categorization or of retrieval from memory as a cue to the validity of the perception or the frequency of the stored event) and thus may misattribute the reason to a plausible and salient possible cause of which he or she is aware.
Consequently, the research literature that is relevant to reviewing the awareness aspect includes stereotyping, construct accessibility, misattribution, perceptual fluency, and subliminality, not to mention mood effects (Ehrlich, 1991; Forgas & Bower, 1987), schematic “capture” effects by knowledge structures sharing representative features with the novel target person or information (Andersen & Cole, 1990; Gilovich, 1981; Lewicki, 1985), and so on. Reviewing the topic of awareness would be a chapter (or a book) in itself; I offer as complete an inventory as possible given the more general purview of this chapter.

Subliminal Perception

Over the past decade, many social psychological studies have demonstrated effects of subliminally presented stimuli. In six of them (Bargh & Pietromonaco, 1982; Bargh et al., 1986; Devine, 1989; Erdley & D'Agostino, 1987; Neuberg, 1988), subliminal trait-related stimuli were presented to activate or prime the corresponding trait concept in memory, making it more accessible and thus more likely to be used subsequently to interpret ambiguous but relevant behaviors (see, e.g., Higgins, 1989; Wyer & Srull, 1986). These trait terms were presented outside of the subject's awareness as part of a first experiment that was allegedly unrelated to the second experiment that followed. Subliminality was achieved by brief presentations of the trait terms, their immediate pattern masking without informing subjects as to the nature of the flashes they saw on the tachistoscope or computer screen, and tests of the subjects' momentary awareness and later recognition memory for the stimuli (see Bargh et al., 1986, for a typical procedure). Bargh and Pietromonaco (1982) found that subjects who were exposed subliminally to hostile-relevant stimuli subsequently rated the target person who behaved in an ambiguously hostile manner as possessing more of that trait than did nonprimed subjects. Bargh et al. (1986) replicated the effect for the traits of kindness and shyness, and demonstrated that such priming combined with the subject's chronic accessibility on these traits in an additive fashion. Erdley and D'Agostino (1987) also demonstrated subliminal priming effects. They also showed (in line with the findings of Higgins et al., 1977) that the effect was not due to a general affective (good vs. bad) priming—the target behavior had to be specifically relevant to the primed construct for the priming effect to occur (i.e., the principle of applicability; see Higgins, 1989).

Devine (1989) used subliminal priming to present elements of the African-American stereotype other than hostility (which pretesting had shown was also an element); subjects primed in this manner rated a subsequent target person ambiguously hostile behavior (Srull & Wyer, 1979) as more hostile than did other subjects. Neuberg (1988) primed subjects subliminally with either competitiveness-related or neutral stimuli, and then had them participate in a Prisoner’s Dilemma game on a computer with a fictitious, preprogrammed partner. For dispositionally competitive but not dispositionally cooperative subjects, the subliminal competitive primes increased the competitiveness of their responses to their partner’s moves throughout the game.

Perdue et al. (1990) demonstrated a context-dependent automatic activation of generally more positive trait concepts when people think about themselves or their in-group than when they think about others or the out-group. After subliminal presentation of in-group primes such as us and we, subjects' response times to trait adjectives in a valence classification task (i.e., “Is this a good or a bad trait for someone to possess?”) were faster for positive than for negative content. However, this effect was (nonsignificantly) reversed when out-group primes such as they and them were the subliminal primes. Dovidio et al. (1986) also demonstrated that different sets of concepts become automatically activated depending on whether one is thinking about one's own group or about those outside that group. Importantly, such differential accessibility of positive versus negative constructs occurs with the merest provocation—priming stimuli such as we or they—suggesting an automatic, cognitive basis for the minimal in-group/out-group effect (e.g., Crocker & Schwartz, 1985; Tajfel, 1970).

Other studies have used subliminal presentation of faces showing a positive or negative emotion to prime affective reactions to a subsequent stimulus (Baldwin, Carrell, & Lopez, 1990; Edwards, 1990; Krosnick, Betz, Jussim, Lynn, & Stephens, 1992; Murphy & Zajonc, 1993; Niedenthal, 1990). Greenwald, Klinger, and Liu (1989) demonstrated subliminal evaluative priming effects in which targets were classified as positive or negative more quickly when a prime of the same valence appeared immediately before it, but more slowly when primes and targets mismatched in valence. Gabriellek and Fazio (1984) showed that subliminal presentation of words containing the letter T resulted in greater frequency estimates for words beginning with that letter, presumably because of the heightened accessibility and ease of recall of memory instances of such words. A separate line of research using subliminal presentation of stimuli studied more exposure effects on liking. In an often replicated finding (Kunst-Wilson & Zajonc, 1980), subliminally presenting some novel stimuli more frequently than others results in subjects' greater liking for those stimuli (Bornstein, Leone, & Galley, 1988; Mandler, Nakamura, & Van Zaandt, 1987; Seamon, Brody, & Kauff, 1983; Seamon, Marsh, & Brody, 1984). Presumably, the buildup of strength in the representation of the more frequently presented stimuli results in their greater ease or fluency of perception (Gordon & Holyoak, 1983), and this in turn results in a positively valenced feeling of familiarity that is misattributed to qualities of the stimulus (because, of course, the subject has no conscious experience of having seen it before). Bornstein et al. (1988) showed that when subjects experienced repeated subliminal exposure to a photograph of a person with whom they had interacted in a group discussion, they agreed with that person's positions more often than with those of a second confederate. There is no longer any doubt, given the abundance of evidence, that environmental stimuli processed outside of awareness can have important interpretive
and evaluative consequences on subsequent conscious thought and behavior (see also Bornstein & Pittman, 1992; Greenwald, 1992). Moreover, recent research has, through technological and methodological improvements, ensured the subliminality of the triggering stimuli more carefully than did the "New Look" subliminal perception research of 40 years ago (see Erdelyi, 1974, for a review). Skeptics of the existence of subconscious processing phenomena have demanded and received rigorous tests, rather than mere claims, of subliminality. Perhaps this increased experimental rigor has done the most to achieve general acceptance of subliminal phenomena (see Bornstein & Pittman, 1992; Greenwald, 1992). Social cognition researchers have been careful to ensure that subjects are not aware of the content or meaning of the subliminally presented stimuli through the use of sensitive forced-choice recognition measures, momentary awareness (as opposed to later memory) tests, and sophisticated pattern-masking procedures (see Bargh et al., 1986, Greenwald et al., 1989, and Niedenthal, 1990, for additional details).

Still, the greater scientific acceptance of the internal validity of the effects obtained in subliminal presentation experiments has not been matched by an acceptance of their ecological validity. If subliminal perception does not happen very often outside of highly artificial laboratory situations that employ specialized procedures and equipment, says the skeptic, why does it matter for social psychology? Why all this subliminality research?

One reason is that subliminal presentation is a methodological tool that researchers have used to ensure that obtained effects were not due to experimental demand or to some other intentional and strategic processing by the subject. This was the motivation behind the use of subliminal priming by Bargh and Pietromonaco (1982). Similarly, researchers of the affect-cognition interface, or of mere exposure effects, have used subliminal presentation of affective primes to demonstrate evaluative reactions to stimuli that cannot be traced to some conscious computation of liking, based on a consideration of the various qualities of the stimulus (see Zajonc, 1980). Devine (1989) used subliminal presentation to show how personal stereotypes can become active and influence judgments without the person's intention to stereotype.

Thus, social psychological research has used subliminal presentation techniques to investigate the ways in which people are not aware of how they interpret stimuli or of the important influences on their judgments. Lack of awareness of the stimulus event ensures that its subsequent effects were unintended by the subject. Thus, subliminality research can be placed in the context of a larger tradition in social psychology, concerning the extent of awareness and control of influences on a person's judgments and behavior (Bargh, 1992a). Cognitive dissonance paradigms counted on the fact that subjects would not be aware of the powerful influence exerted on their free choice by the experimenter, and so would attribute their decisions to some internal factor (e.g., Festinger & Carlsmith, 1959; Wicklund & Brehm, 1976). Such a lack of access to the causal influences on people's decisions was seized on by self-perception theorists (e.g., Bernt, 1972) as a different explanation for dissonance findings. Other attribution research has discovered many other such misattributational tendencies, including those involved in understanding the source of one's internal states such as emotions, arousal, or mood (e.g., Schachter & Singer, 1962; Schwarz & Clore, 1983; Zillman & Bryant, 1974), or one's social and nonsocial judgments (e.g., Nisbett & Bellows, 1977; Nisbett & Wilson, 1977). More recent subliminal presentation studies are direct descendants of this research tradition into the hidden influences on phenomenal experience and social judgment.

In other words, social psychologists have not been studying subliminality per se, but have used subliminal presentation as a tool to study how people can be unaware of many important but unintentional influences over their judgments and behavior. Subjects who are unaware of the stimulus that causes an effect obviously do not intend for the effect to happen, and consequently they are unable to control the effect (Bargh, 1988; Devine, 1989; Fiske, 1989).

Awareness of the Stimulus Versus Awareness of Its Influence. Perhaps the most important reason why subliminality is not of prime concern for social psychology is that similar results are obtained with supraliminal stimulus presentation as long as subjects are not aware of the influence of that stimulus. Studies using conscious presentation of the critical stimuli have repeatedly produced the same findings as studies using subliminal presentation, providing the relation between those stimuli and subsequent processing tasks has been obscured. Bargh and Pietromonaco (1982), Bargh et al. (1986), Erdley and D'Agostino (1987), and Devine (1989) all obtained assimilative priming effects with subliminal presentation of the primes—the same effect obtained in conscious priming studies when subjects are unaware of the possible influence on their subsequent judgments (see Higgins & King, 1981). Several studies have shown affective reactions to neutral stimuli in line with subliminally presented emotional faces (e.g., Edwards, 1990; Krosnick et al., 1992; Murphy & Zajonc, 1993; Niedenthal, 1990), but other studies have shown similar effects of affect-inducing stimuli of which subjects were consciously aware but did not realize the potential effect (e.g., eye-pupil dilation: Niedenthal & Cantor, 1986; incidental touch: Craske & Wetzel, 1984). For example, a brief incidental touch by a waitress when returning change increased the size of the tip she received (Craske & Wetzel, 1984),

2This is not to say that it does not matter at all, the existence of subliminal phenomena obviously matter for questions such as the nature of consciousness, psychodynamic influences, and its potential for misuse (such as by advertising, governments, etc., see reviews in Bornstein & Pittman, 1992).

3Although, as argued here, the quality of the effect is the same for subliminal and supraliminal presentation, one might suspect that the size of the effect would be greater for stimuli presented supraliminarily, given that they impinge on the senses longer and are of greater intensity (that is what makes them supraliminarily after all). Although this logic holds for category priming effects, there are nonetheless domains in which subliminal effects are the stronger (e.g., in mere exposure effects; see Bornstein, 1989).
and similar behavior by a librarian when returning a library card resulted in subsequent more positive ratings of the library (Goleman, 1988).

Baldwin and Holmes (1987) showed that prior conscious exposure to a significant other affected subjects’ evaluations of themselves, and Baldwin et al. (1990) obtained the same effect using subliminal presentations of the faces of significant others (the Pope to observing Catholics and Bob Zajonc to observing Michigan graduate students). Mere exposure effects of greater liking occurred with both conscious and incidental exposure of the novel stimuli (Moreland & Zajonc, 1977; Zajonc, 1968) and with subliminal presentations (see review by Bornstein, 1989).

Devine (1989, Experiment 2) produced stereotype influences on judgments using stereotype-relevant subliminal priming words, whereas the same stereotype influences have been produced by consciously perceived target persons or information (e.g., Darley & Gross, 1983; McArthur & Friedman, 1980; Pratto & Bargh, 1991; Rosenfield, Greenberg, Foiger, & Borys, 1982; Sagar & Schofield, 1980). In one experiment, Edwards (1990) subliminally presented a positive or negative facial expression as a prime before exposing subjects to the target-attitude objects. In a second experiment, she presented those faces at supraliminal durations. The manipulation induced an affect-based attitude toward the attitude object in both studies, regardless of whether the prime was in or out of awareness.

As a final example, the Greenwald et al. (1989) experiment described earlier found evaluative priming effects for a subliminally presented prime. The same effect has been found repeatedly for supraliminally, but briefly (250 msec), presented evaluative primes in the same evaluative judgment task (Barth et al., 1992; Fazio et al., 1986).

It is clear from these findings that awareness of the stimulus does not matter to an effect as long as subjects are unaware of the potential influence of that stimulus. When subjects are aware of that potential influence, different effects occur. In category accessibility studies, in which the priming stimuli are still in working memory at the time of the subsequent impression formation task so that subjects could be aware of the potential influence of the priming events on their judgments, contrast rather than assimilation effects are often obtained. The likelihood of the priming stimuli continuing to reside in working memory has been manipulated by the extremity or vividness of the primes (e.g., Dracula as a hostile prime; Herr et al., 1984), by interruption of the priming task (Martin, 1986), and by the subjects processing the priming stimuli with greater effort (Martin, Sela, & Crelia, 1990). Alternatively, Lombardi, Higgins, and Bargh (1987) and Newman and Uleman (1990) assessed whether subjects could recall the primes at the time of the impression-formation task. In all these studies, the residence of the primes in consciousness produced contrast effects in judgment, instead of the assimilative effects obtained when subjects were unaware of the potential influence of the primes.

Other research domains show the same critical role for awareness of a potential influence as opposed to awareness of the critical stimulus. In stereotyping research, in which the subjects are aware of the possible influence of stereotypes on their judgments and descriptions and are motivated to control these influences, they can do so (Devine, 1989, Experiment 3; Pratto & Bargh, 1991). In the Pratto and Bargh (1991) pretesting, between-subjects tests of the existence of gender stereotypes—in which a subject rated either the average male or the average female—were successful in documenting the stereotypes obtained in much previous research (e.g., Ruble & Ruble, 1982), whereas within-subjects tests—in which the same subject rated both the average male and the average female on the same traits—showed no stereotyping at all.

Summary. Subliminal research, then, is important for understanding what kinds of effects occur naturally, immediately, and unintentionally on the part of the subject. What is critical for whether the effect occurs is not subliminality itself but the subject’s awareness of the possibility of the influence by that stimulus as well as the subject’s values and motivations (see Controllability section) to control that influence.

In this regard, it is important to distinguish between a person’s awareness of a stereotype or an accessible construct, and the actual influence of the construct on the person’s judgment. One cannot be aware of the actual occurrence of accessibility or stereotypic influences because of the fast, effortless, and immediate (i.e., preconscious) way in which those mental structures capture and interpret relevant environmental input. Nonetheless, through education and other consciousness-raising techniques, one can become aware that one might be influenced. For example, one may have no conscious experience of stereotyping an African-American assistant professor applicant, a female engineering graduate school applicant, or a Korean colleague, but might nonetheless take steps in reporting one’s judgments and decisions to adjust or counteract these potential influences of the stereotypes. For instance, one could perform a more deliberate and effortful conscious appraisal of the individuating qualities of the person than one would normally (see Fiske, 1989; Fiske & Neuberg, 1990; Thompson, Roman, Moskovitz, Chaiken, & Bargh, 1992). If that is not possible, one might adjust one’s opinion somewhat in the direction opposite to the assumed stereotypic influence (Strack, 1992) or consider the opposite conclusion (Lord, Lepper, & Preston, 1984).

Misattribution

As argued previously, a lack of awareness of an influence on thought or behavior matters, because it precludes the possibility of controlling that influence. Another way such unawareness of influence matters is that one might misattribute the cause or source of one’s impressions of another or one’s own subjective state to more salient potential causes (see Nisbett & Wilson, 1977).

Schwarz and his colleagues (e.g., Schwarz, 1990, Schwarz & Clore, 1983)
demonstrated that people often are unaware of the reasons for their current mood (for example, whether it is a sunny or a rainy day) and, unless these true causes are called to their attention in some way, will attribute those moods to whatever is currently salient in their environment—even to a general satisfaction or dissatisfaction with their life if they are being asked to complete such a questionnaire (Schwarz & Clore, 1983).

As Schwarz (1990) argued, current feeling states can serve as a source of information for an individual in making decisions and judgments when the source of those feelings is assumed, correctly or incorrectly, to be the person or topic being judged. Another kind of subjective feeling that has been studied for its non-conscious and misattributational effects is the feeling of ease or fluency in perception that comes from prior experience (Bargh, 1992a; Jacoby, Toth, Lindsay, & Debner, 1992; Smith, chapter 3, Volume 1). Usually, the felt ease of categorizing or perceiving a person or event is a diagnostic cue to the validity of that categorization, either because the person or event unambiguously matches the features of that category or because of the frequency and consistency of mapping that person or event to the category in the past (see Fiske & Neuberg, 1990; Higgins & Bargh, 1987). For example, Anita’s victory in the university chess championship is effortlessly understood as an intelligent act. Here the bottom-up strength or diagnosticity of the behavior determines the ease of comprehension; there is no need to engage in a “search after meaning” (see Postman, 1951).

However, suppose that one has a strong expectancy about an individual, a strong stereotype about a group, or a chronically accessible construct concerning people in general. These top-down influences of accessibility can also result in a subjective feeling of ease or effortlessness in perceptual categorization, or perceptual fluency, even with relatively ambiguous and nondiagnostic input (see Higgins, 1989). The consequence is that people often misattribute the source of the fluency caused by the top-down expectancy or accessibility to the diagnosticity of the stimulus. People are quite aware of the stimulus person or behavior; they are less aware of the effect that their own readiness to perceive the person or behavior in certain terms has on the ease of doing so.

**Construct Accessibility as Perceptual Fluency.** Recently, several authors argued that such accessibility or readiness effects can be conceptualized as perceptual fluency effects (Bargh, 1992a; Schwarz, Bless, Strack, Klump, & Simons, 1991; Sherman, Mackie, & Driscoll, 1990; Smith & Branscombe, 1987, 1988; Spielman & Bargh, 1990). For example, Sherman et al. (1990) primed certain dimensions that were relevant to judgments of a target politician’s abilities in either foreign affairs or managing the economy. They hypothesized that subjects would attribute the greater ease of perceiving and categorizing the information with respect to the primed dimensions (relative to the unprimed dimensions) to the validity or diagnosticity of the information. Consistent with the hypothesis, dimensions were given greater weight in subjects’ overall judgments. A similar effect of chronically accessible constructs on the weight given by various behaviors in overall liking judgments was obtained by Spielman and Bargh (1990).

Schwarz et al. (1991) also showed that the felt ease of retrieval from memory of relevant information is taken as a cue in memory-based judgments. Although this is the same logic as that of the availability heuristic (Kahneman & Tversky, 1973), it had never been tested directly with an experimental manipulation of felt ease of retrieval while holding the amount of retrieval constant. Schwarz et al. (1991) accomplished this by asking subjects to recall either 8 or 12 instances of times when they behaved in a given trait-like fashion, following which subjects rated themselves on that trait dimension. Pretesting had shown that subjects were able to come up with 8 examples much more easily than 12 examples, so that subjects in the recall-12 condition would experience greater difficulty completing the task than the other subjects. Results showed that although the recall-12 subjects remembered more examples of that trait than did subjects in the recall-8 condition, they nevertheless rated themselves as possessing less of that trait than did the other subjects, in line with the retrieval fluency as cue hypothesis.

Jacoby, Kelley, Brown, and Jaseckho (1989) manipulated perceptual fluency by exposing subjects to a series of nonfamous along with famous names. The next day, the previously exposed nonfamous names were more likely than completely novel names to be mistaken as famous. Again, the feeling of familiarity that subjects presumably felt while seeing the name again was misattributed to the fame of the name.

**Conclusions**

Awareness as an aspect of automaticity is a critical issue for the intentional control of thought and behavior. What matters more than whether one is aware of a stimulus event is whether one is aware of the potential influence of that event on subsequent experience and judgments. All sorts of influences exist of which one does not have conscious knowledge, from immediate and unintended affective reactions to current moods to subjective feelings of familiarity and perceptual fluency. Thus, one attributes these effects to those environmental features one does have conscious knowledge of and that seem plausible causes of one's reactions. This phenomenon was described over 20 years ago when Jones and Nisbett stated, “[O]ne tends to regard one's reactions to entities as based on accurate perceptions of them. Rather than humbly regarding our impressions of the world as interpretations of it, we see them as understandings or correct apprehensions of it. . . . The distinction between evaluations and primary qualities is never fully made. We never quite get over our initial belief that funniness is a property of the clown and beauty is in the object” (1971, p. 86).
INTENTIONALITY

The intentionality and controllability aspects of automaticity both have to do with how much one is in control of one's own thought and behavior. Intentionality has to do with whether one is in control over the instigation or "start up" of processes, whereas controllability has to do with one's ability to stifle or stop a process once started, or at least to override its influence if so desired. To the extent that perceptual, judgmental, and behavioral processes are triggered by the environment and start up without intention, the environment is more in control (see Bargh & Gollwitzer, in press). To the extent that these processes, once started, can be stopped by an act of will, they are controllable by the individual (see Logan & Cowan, 1984).

Automatic Attention Responses and Perceptual Selection

Two kinds of automatic attention responses (Shiffrin & Schneider, 1977) have been documented in social cognition: (a) responses to information relevant to accessible trait constructs and attitudes, and (b) responses to negatively valenced stimuli. Behaviors clearly relevant to a person's chronically and temporarily accessible trait constructs are more likely to receive attention and be remembered later (Higgins, King, & Mavin, 1982; Sherman et al., 1990), to be noticed and influential in impressions even when attentional processing is severely constrained (Bargh & Thein, 1985), and to draw attention even when the subject is trying to ignore them in a dichotic listening task (Bargh, 1982) or a Stroop color-naming task (Bargh & Pratto, 1986; Higgins, Van Hook, & Dorfman, 1988). For example, Bargh and Pratto (1986) found that subjects took longer to name the color of trait terms corresponding to their chronically accessible than their inaccessible trait constructs. Recently, Roskos-Ewoldsen and Fazio (1992) obtained the same uncontrollable distraction effect for attitude objects when their associated attitude is made temporarily more accessible through its repeated expression by the subject. Therefore, the greater the accessibility in memory, the less subjects are able to prevent devoting processing resources to the corresponding behaviors or attitude objects. As a result, behaviors and objects are more likely to be noticed and be influential in on-line judgments and behavioral decisions.

A second determinant of automatic attention responses is negative social stimuli, in terms of either undesirable behavior (Fiske, 1980), negatively valenced trait terms (Pratto & John, 1991), or faces expressing negative emotions (Hansen & Hansen, 1988). The latter study applied Shiffrin and Schneider's (1977) method of varying the size of the stimulus array through which subjects had to scan to find an angry or happy face. Angry faces seemed to pop out of an array of happy ones; that is, subjects were able to respond quickly when asked whether an angry face was present, and increasing the number of distractor faces did not increase response time (as it would if subjects were engaging in an attentional, serial search process; see Shiffrin & Schneider, 1977).

Pratto and John (1991) used the Stroop task to show longer color-naming latencies (i.e., more uncontrollable distraction) for undesirable than for desirable trait terms, and greater incidental recall of the undesirable trait terms as well. Subsequent experiments ruled out possible artifactual explanations in terms of differences between the desirable and undesirable trait terms in their length, frequency, or the perceived base rates of occurrence for corresponding behaviors. The uncontrollability of this attention response was demonstrated in further studies reported by Pratto (in press), in which the negative trait concepts caused greater distraction even when subjects were informed of the effect and exhorted to overcome it. Pratto and John (1991) couched their predictions in the context of a model of automatic vigilance, in which attention is automatically given to stimuli and events that might affect the individual negatively. Taylor (1991) also has described an immediate mobilization response in the face of negatively impacting events.

Automatic Evaluation

To be able to immediately notice and attend to negative events, it is necessary to posit an earlier stage of processing in which all incoming stimuli are classified as positive or negative. The results of several recent studies have been consistent with this immediate classification or evaluation stage. Bargh, Litt, Pratto, and Spielman (1989) conducted a replication and extension of Marcel's (1983) study of preconscious analysis of meaning. In that study, a subject answered questions about words presented tachistoscopically for durations that were below his or her individually established recognition threshold. The questions concerned whether a word had been presented at all, the physical characteristics of the word (i.e., whether it was presented in upper- or lower-case letters), and the semantic meaning of the word (i.e., whether another word was a synonym of the target word). Marcel (1983) found that subjects responded at better than chance levels about the semantic meaning of words at presentation durations at which they could not answer the other questions at more than random guessing levels.

Bargh et al. (1989) used trait words as stimuli and added an evaluative question to Marcel's basic design. That is, on any given trial, subjects answered the presence or semantic question about the subliminally presented word, or responded as to whether the word was positive or negative in meaning. Bargh et al. selected the stimuli from Anderson's (1968) normative ratings of traits as to their likability, choosing sets of moderate and extreme and positive and negative adjectives. On successive blocks of trials, words were presented at faster and faster durations.

As predicted, subjects were able to answer the evaluative question at better than chance levels for presentation durations in which they could not answer the semantic question at nonrandom levels. Path analyses confirmed that the subjects' ability to answer the evaluative question correctly was statistically
independent of their ability to answer the semantic question correctly. Moreover, the extremity or intensity of the adjectives' evaluative meaning did not matter to these effects. What did matter was whether the stimulus was positive or negative in valence, regardless of its extremity. In other words, subjects had access to the polarity of the trait adjective's valence in the absence of access to other aspects of its meaning, and this knowledge was independent of the extremity of this valence.

This dichotomous preconscious classification of stimuli by valence recalls Neisser's (1967) argument that such preconscious analyses of environmental stimuli are crude and basic, not fine-grained. It also supports the argument of Swann, Hixon, Stein-Seroussi, and Gilbert (1990) concerning the priority of self-enhancement over self-verification responses to self-relevant feedback. These authors posited an initial immediate classification of the feedback as favorable or unfavorable, followed by attention-demanding self-verification only if sufficient resources were available.

Research on the automatic activation of attitudes has also led to the conclusion that there is an initial automatic evaluative classification of stimuli as good or bad that does not vary as a function of the intensity or extremity of the stimulus valence. Fazio et al. (1986) found that a subject's relatively strong attitudes, not his or her relatively weak ones (defined in terms of how quickly subjects could evaluate the attitude object in a previous assessment task), were capable of becoming active automatically in the context of an adjective evaluation task. Attitude objects selected from the assessment phase of the study were employed as priming stimuli in the adjective evaluation task, being presented too briefly (250 msec) to permit intentional, strategic evaluation to occur (see Neely, 1977; Posner & Snyder, 1975). On each trial, one of these strong or weak, good or bad attitude object primes was presented, followed by a target adjective that was clearly positive (e.g., beautiful) or negative (e.g., repulsive) in meaning. Subjects were to classify each adjective, as quickly as possible, as having a positive or negative meaning, by pressing either a “good” or a “bad” button. When the attitude object primes corresponded to the subject's strongly held attitudes, responses were faster when the prime and target evaluations matched than when they mismatched. The effect when primes corresponded to weak attitudes was less evident. Thus, even though subjects were asked to evaluate the adjective targets and not the primes themselves, the strong-attitude primes apparently activated their stored evaluation and consequently facilitated or interfered with evaluating the adjectives with which they were paired. Fazio et al. (1986) concluded that one's strong, relatively accessible attitudes become active automatically at the mere presence of the attitude object in the environment.

Fazio et al. (1986) concluded that the mere presence of the attitude object in the environment was sufficient to activate its associated attitude and, therefore, to influence on-line judgment and behavior concerning the object. However, several aspects of the paradigm they used to assess preconscious automaticity potentially could have activated the attitude through postconscious or goal-dependent means instead. Specifically, subjects were instructed to think about and give their attitudes for each possible prime immediately before testing the automaticity of those attitudes. This procedural step could have increased the temporary accessibility of the relevant attitudes, producing a postconscious automaticity effect that requires recent conscious thought about the attitude object. Moreover, because subjects were intentionally and consciously evaluating the target adjectives while the attitude object primes were being presented, it is possible that the evaluation of the primes depended on subjects having the evaluative processing goal at the time the primes were presented (see Gollwitzer, Heckhausen, & Steller, 1990; Mandler & Nakamura, 1987).

My colleagues and I (Bargh et al., 1992; Chaiken & Bargh, 1993) found that when the original paradigm was altered to eliminate the possibility of postconscious and goal-dependent activation of attitudes, the automatic activation effect was obtained for all attitude objects, regardless of their relative strengths or accessibilities. For example, in two experiments we inserted a 2-day delay between the attitude assessment phase of the experiment and the adjective evaluation task that assessed automaticity. Because subjects had not evaluated the target consciously, the attitude would not be temporarily more accessible in memory. Nonetheless, the effect identified by Fazio et al. (1986) was maintained. It occurred more generally than it had in the Fazio et al. (1986) studies, however, with even the weakest (i.e., most slowly evaluated) of the subjects' attitudes from among the range of stimuli presented showing the effect.

More recently, we examined the possible goal dependence of the effect by eliminating the adjective evaluation task from the paradigm (Bargh, Chaiken, Raymond, & Hymes, 1993). Specifically, we had subjects pronounce the adjective targets as quickly as possible, and assessed how quickly they could do so under the various prime valence × target valence combinations, as before. Removing this potential condition for the effect and making the experimental situation even more like conditions of mere presence of the attitude object did not eliminate the effect. Rather, the effect again occurred for all attitude object primes, regardless of whether they corresponded to the subject’s strongest or weakest attitudes, and was of equivalent strength across the range of attitude strengths. Therefore, under conditions more closely resembling the mere presence of the attitude object in the environment, it appears that nearly everything is preconsciously classified as good or bad, with this effect occurring equally strongly regardless of variations in the underlying strength of the attitude; that is, in the “crude” dichotomous manner demonstrated in the Bargh et al. (1989) and Pratto and John (1991) studies discussed previously.

4This more general automatic evaluation effect recently has been obtained with complex pictorial stimuli as primes and targets as well as with word stimuli (Giner-Sorolla, Chaiken, Bargh, & Garcia, 1993; Hermans, de Houwer, & Eelen, 1992), so it would appear not to be merely a verbal effect.
A Methodological Caution. There are important methodological consequences of the existence and ubiquity of this preconscious evaluation effect. The Fazio et al. (1986) and Bargh et al. (1992, 1993) demonstrations of automatic attitude activation used primes and targets that were matched or mismatched randomly on valence alone—they had no other semantic features in common (see also Greenwald et al., 1989). The priming stimuli somehow must have activated all similarly evaluated material in memory, making it immediately and, for at least a short time, more accessible than opposite-valence material in general (see Bargh et al., 1993, for a fuller discussion of mechanism).

Hence, the results of other sequential or context-dependent priming studies that were interpreted in terms of specific features of the primes may have occurred because of correlated differences in the valence of the primes. For example, in a study of age stereotyping by Perdue and Gurzman (1990, Experiment 2), subjects on each trial were subliminally presented with the word young or the word old, followed immediately by a positive or negative adjective they were to classify as good or bad, following the Fazio et al. (1986) paradigm. Subjects were faster to respond to positive adjectives following young and to negative adjectives following old, and this was interpreted in terms of the automatic activation of a positive stereotype of young people and a negative stereotype of older people by subjects (who were college students). However, it is likely that young is positive and old is negative in meaning. Greenwald et al. (1989) showed such priming effects of subliminally presented stimuli based only on the valence match or mismatch between prime and target.

Summary. Collectively, the evidence in this domain indicates that the automatic, preconscious evaluation of stimuli is a ubiquitous and constant mental process. It leads input to be classified immediately as good or bad, regardless of the intensity, extremity, or strength of that evaluation or affective reaction. At least this is what occurs unconditionally, upon the mere presence of the stimulus in the environment. Following this initial preconscious screening of the environment, there may be differential processing of stimuli based on their self-relevance (e.g., Lazarus, 1991), attitude strength (e.g., Roskos-Ewoldsen & Fazio, 1992), or survival implications (e.g., Pratto & John, 1991). Certainly, the results of the Fazio et al. (1986), Bargh et al. (1992, Experiments 1 and 3), and Roskos-Ewoldsen and Fazio (1992) studies showed variations in the size of the automatic evaluation effect with differences in underlying attitude strength when one has recently thought about one's attitude toward the object. Thus, just because there are no differences in the size or extent of the unconditional automatic evaluation effect does not mean that such differences do not occur given certain conditions.

The ramifications are considerable for a preconscious evaluative process that immediately classifies everything and everyone the individual encounters as either good or bad, because of its potential influence on subsequent judgments (e.g., how one interprets a person's ambiguous behavior) and behavior toward the person or object. The importance of immediate affective reactions for subsequent cognitive processing has already been noted by theorists such as Niedenthal (1990) and Niedenthal and Cantor (1986). Given the automatic evaluation evidence, such reactions may be a more pervasive and constant influence than was previously assumed.

Automatic Stereotype Activation

It has been argued widely that stereotypes are activated automatically by the presence of a group member, as easily identified by physical characteristics such as skin color or gender features, or by accent, dress, and so on (Brewer, 1988; Deaux & Lewis, 1984; Devine, 1989; Fiske, 1989; Perdue & Gurzman, 1990; Pratto & Bargh, 1991; Rothbart, 1981). This activation appears to be unintentional and efficient, at least for the more widely shared stereotypes (within the U.S. culture), such as those for African-Americans and for different genders. Devine (1989, Experiment 2) subliminally presented subjects with words related to the African-American stereotype, both positive (e.g., musical) and negative (e.g., lazy), but none related to hostility, which her Experiment 1 had shown to be part of the stereotype. Subjects then read about a target person (race unspecified) who acted in an ambiguously hostile manner. Subjects who were primed with the stereotype-related words rated the target person as more hostile than did control subjects. Apparently, the African-American stereotype was activated by the prime words and caused the unprimed trait concept of hostility to become activated and more accessible by virtue of its inclusion in that stereotypic representation (i.e., all-or-none activation; see Fiske & Dyer, 1985; Hayes-Roth, 1977). Devine's (1989) set of studies was ground breaking conceptually, because of its analysis of the stereotyping process into separate components of stereotype activation and stereotype use, and empirically because it demonstrated the relative controllability of the latter but not the former stage (see next section). However, as a single article, it could not be expected to address and answer each question having to do with stereotype activation. There are intriguing aspects of the findings that call for further study, especially as to the inevitability of stereotype activation in more natural settings.

Most important of these is that the race of the target person whom subjects rated after the subliminal priming task was not specified in the story (subjects read the Donald story used by Srull & Wyer, 1979, in which Donald behaved in ambiguously hostile ways on several occasions). Presumably, most subjects assumed that Donald was White, given base rates and the fact that all subjects in the study were White. In effect, then, the real-world analogue to the results...
of Experiment 2 would be if the mere presence of an African-American in the current environment caused the perceiver to categorize a White's (or anyone's) ambiguously hostile behavior as more hostile than would a perceiver who had not just encountered an African-American. However, the general assumption about the application of group stereotypes is that they are used in interpreting (or making assumptions about) the behavior of group members, rather than nongroup members who happen to be in their vicinity. Thus, although Devine's (1989, Experiment 2) results were suggestive and provocative, they signaled the need for further research to better understand their implications for automatic stereotype activation and application.

Such additional study is needed all the more in the wake of a recent experiment by Gilbert and Hixon (1991). Subjects watched a videotape in which an Asian-American experimenter held up word-fragment completion items for subjects to complete. Five of these were critical trials (e.g., S __ Y) that had stereotypic (e.g., SHY) as well as nonstereotypic completions (e.g., STY, SPY). The stereotypically was determined by pretest assessment of the Asian-American stereotype among the subject population. With no constraints on attentional capacity, the incidental presence of the Asian-American experimenter did result in a greater number of stereotypic completions compared to the Caucasian experimenter condition. However, in two experiments, giving subjects a simultaneous digit-recall task to constrain attentional processing eliminated the stereotyping effect. Apparently, then, at least for some stereotypes, activation is unintentional, but requires attentional capacity. Further research is needed to determine whether this holds true for other, perhaps more strongly held stereotypic beliefs (as for women or African-Americans). At a minimum, such findings do question the assumption that stereotype activation is inevitable.

Such provisos notwithstanding, Gilbert and Hixon (1991) made an excellent point when attempting to reconcile their findings with those of others (Devine, 1989; Perdue & Gurtman, 1990; Pratto & Bargh, 1991) who have concluded that group stereotypes are automatically activated given the presence of features of a group member. These contrasting findings come from experiments in which the stereotype was primed or activated using verbal labels or descriptions that may force a categorization in terms of the stereotyped group, whereas an actual person displays many other features (height, age, expensiveness of dress, self-confidence, accent, etc.) besides race, gender, or ethnic group membership that also can be used to categorize the person (see also Zarate & Smith, 1990).

For example, if, in an experiment, the subject is told only that the target is elderly, he or she may assume implicitly that the target is passive, needy, and physically weak (see Perdue & Gurtman, 1990); if the subject is told only that the target is an African-American male, he or she may assume implicitly that the target is hostile, athletic, and aggressive (see Devine, 1989). Does this mean that all of these trait expectations are activated automatically in the presence of an elderly African-American? They would seem to be mutually contradictory.

It may be that people have more specific subtypes that become activated automatically (e.g., Taylor, 1981; Weber & Crocker, 1983), or that stronger stereotypes override weaker ones (e.g., the elderly stereotype overrides those for minority group membership). Consistent with this reasoning, Brewer and Lui (1989) examined the priority with which identifying features are used in categorizing people, and found that age and gender are the paramount determinants. Such results call for a more specific and conditional model of automatic stereotype activation than currently exists.

Proceeding down the road suggested by Gilbert and Hixon (1991) and Zarate and Smith (1990; see also Smith & Zarate, 1992), it seems useful to consider real people as collections and combinations of features instead of existing as placements on single dimensions. Thus, stereotypes may not exist at a global abstract level, but rather for specific, concrete exemplars or instances of people with certain combinations of features. For example, instead of a single stereotype triggered by group membership regardless of other features (e.g., African-American woman), it may require multiple features to become active in the natural environment (e.g., young, male, poorly dressed African-American; middle-aged White female).

Spontaneous Trait Inference

If there is one social-cognitive process that is automatic in all senses of the word, it is the identification or categorization of social behavior in trait terms when that behavior is diagnostic of a trait (i.e., unambiguously relevant to the trait construct; see Higgins, 1989). In their study of priming effects on impression formation, Srull and Wyer (1979) assumed this automatic behavior-to-trait process when they used short sentences indicating hostile or kind behaviors as the priming stimuli. They presented these behavioral examples in scrambled-sentence form (e.g., "the kick he dog"), with subjects instructed simply to make grammatical three-word sentences out of the word string. Although the ostensible purpose of this experiment (i.e., to measure language ability) had nothing to do with personality or impression formation, these behaviors nonetheless primed the corresponding abstract trait construct. In the subsequent, "unrelated" second experiment, subjects formed impressions of a target person whose behavior was ambiguously relevant to the primed trait, and primed subjects considered the target to possess more of that trait than did nonprimed subjects. More recently, Moskowitz and Roman (1992) also showed that trait-implying behavior descriptions have this priming function, although subjects are instructed only to memorize the sentences. Thus, at least with verbal presentation of the behavioral stimuli, behaviors activate corresponding trait concepts unintentionally and without subjects' awareness of such encoding (i.e., "spontaneously"; see Newman, 1991; Newman & Uleman, 1989).

Winter and Uleman (1984) and Winter, Uleman, and Cuniff (1985) used an
a given trait term was true of the self or of the average other person. Depressives made self-judgments on depressed-content traits just as quickly under the load as under the no-load conditions, whereas nondepressed subjects did the same for the nondepressed-content trait, supporting the hypothesis that, when thinking about the self, different content becomes active automatically for depressed versus nondepressed people (both groups of subjects thought about other people most efficiently in terms of nondepressed constructs). Recently, Andersen, Spielman, and Bargh (1992), using the same memory load technique, showed depressed subjects to respond automatically to questions about the likelihood of future events in their lives.

Smith and Lerner (1986; see also Smith, chapter 3, Volume 1) used a response-time measure to show how subjects given the task of judging whether behaviors are instances of specific traits make these judgments more efficiently (faster) with practice, with this procedural knowledge having both specific behavior-to-trait (Smith, Stewart, & Buttram, 1992) and more general skill components (Smith, Branscombe, & Bormann, 1988).

The ways one thinks about oneself or others under attentional stress, and the kinds of information that are picked up about others regardless of concurrent attentional focus or demands, are quite important, because such processes operate (given the goal to do so) much more routinely than do processes that are dependent on the current availability of sufficient attentional capacity for their occurrence.

Dispositional Inference

Winter et al. (1985), Lupfer et al. (1990), and Uleman, Newman, and Winter (1992) examined the efficiency of spontaneous trait inferences using a concurrent memory load technique. Whereas Winter et al. (1985) and Lupfer et al. (1990) found that their secondary task (digit retention) did not interfere with spontaneous trait inferences, indicating their efficiency, Uleman et al. (1992) added a probe reaction-time measure of spare processing capacity, and did obtain interference. Perhaps the Uleman et al. (1992) probe reaction-time task, when added to the other secondary task of digit retention, constituted a greater attentional load than experienced by subjects in the previous two experiments. Thus, it appears that the spontaneous, unintentional encoding of behaviors in trait terms (see Intentionality section) is at least a somewhat efficient process as well.

Gilbert and his colleagues performed a legion of demonstrations of the effect that attention load, or cognitive busyness (Gilbert & Osborne, 1989), has on causal attribution processes (Gilbert et al., 1988; Gilbert & Osborne, 1989; Gilbert & Krull, 1988). Gilbert posited a three-stage process of (intentional) person perception: an immediate characterization of behavior in trait terms, a dispositional inference stage, followed by a correction stage in which situational reasons for (or constraints on) the behavior are taken into account. Thus, Gilbert's explanation for the correspondence bias or fundamental attribution error was that dispositional attributions are made first and with great ease and efficiency, with situational attributions possible only if sufficient time and attention are available to the perceiver.

For example, Gilbert et al. (1988) showed subjects a videotape of a woman who was said to be discussing either an intimate, embarrassing topic (e.g., sexual fantasies) or a mundane topic (e.g., hobbies). Thus, there was either a situational reason or not for her somewhat anxious appearance (subjects only saw and did not hear the woman on tape). Some subjects were given a secondary task to load attention while watching the tape, and others did not have this constraint on processing resources. The former group of subjects considered the woman to be more dispositionally anxious than did the nonoverload subjects. Thus, even though all subjects had both the relevant dispositional and situational information available to them, the capacity-limited subjects were unable to use the situational information or to integrate it with the dispositional information to adjust the more efficiently made dispositional inference (see Pratto & Bargh, 1991, for a related finding).

Gilbert (1991) placed this efficient dispositional inference phenomenon in the larger context of a general tendency for people to initially believe or accept propositions as true. This belief or acceptance is said to occur naturally during the process of comprehending the meaning of the incoming information, and only subsequently do people correct or adjust this primary trust in the face of reasons to believe otherwise (e.g., one's own knowledge or experience, the possible motives of the source of the information). If dispositional attributions are made naturally and efficiently in the course of one's attempt to comprehend the meaning of another's behavior, then they too will be accepted as valid if the effortful situational-correction stage is prevented in some way (no time, too much to attend to, etc.).

Conclusions

Perhaps all of these efficient trait categorizations and attributions described in this section are trusted precisely because of their efficiency, in that people experience them as being made effortlessly, as conclusions reached easily (see previous discussion of the use of perceptual fluency as a cue for the validity of the conclusions).
inference). If so, this is another reason why the relative efficiency of a mental process matters. People’s trust in the validity of efficiently reached categorizations, self-judgments, future judgments, and attributions may be of the same cloth as their necessary trust in what their senses are telling them, which also comes to them, not coincidentally, with a subjective experience of effortlessness (see Bargh, 1989).

CONTROLLABILITY

There has been a surge of attention given to studying how the subject’s motivations can moderate or even eliminate otherwise automatic (unintended, efficient, unaware) influences on judgments and behavior. As with the Awareness section, it is useful here at the outset to call for some precision in describing what exactly is being controlled in these studies. For example, Devine (1989) was careful to distinguish the process of stereotype activation from that of making stereotypic judgments, and both Trope (1986) and Gilbert (1989) distinguished between a behavior identification and a situational correction stage in their attribution models. In all three of these approaches, the first stage is seen as much less easily controlled than the second. Therefore, such distinctions are important for any discussion of controllability of thought and behavior, because they demonstrate that asking whether stereotyping or dispositional attributions occur automatically are meaningless questions. Just as with other complex mental phenomena, such as those involved in driving a car, social cognitive processes are composed of both automatic and controlled subprocesses.

Thus, what most researchers mean by the question of controllability is not the occurrence of the stereotype’s or accessible construct’s input into a judgment, but rather whether one is aware of such influences and is both motivated and able to counteract them. In an engaging treatment of this issue of ultimate control, Fiske (1989) argued that it is possible to gain control by “making the hard choice” and spending the additional cognitive effort to avoid pigeonholing or stereotyping an individual. Instead, the person can effortlessly seek out additional individuating information and integrate it into a coherent impression (see also Brewer, 1988; Fiske & Neuberg, 1990). It may be that all one can do with this extra effort is to adjust one’s judgment in the direction opposite to that of the suspected stereotypic influence (see Bargh, 1992a, 1992b; Martin et al., 1990; Schwarz & Bless, 1992; Strack, 1992), but doing so is still an act of control.

Under what conditions will a person go this extra mile? If one processes information about the target person more effortfully, even if there are stereotypic or categorical inputs into one’s judgments (as when an influence exists that the perceiver is not aware of and therefore does not engage in an adjustment process; see Bargh, 1989), those judgments will at least be moderated by the additional individuating information collected, and will not be determined solely by the stereotypic input (Fiske & Neuberg, 1990). Many such motivations have now been documented.

Situationally Induced Motivations

Lord et al. (1984) showed that confirmatory biases in hypothesis testing can be overcome by simply instructing subjects to consider alternatives, or the possibility that the opposite conclusion could be correct instead. One will also be more likely to process information about another individual effortfully when that individual has power or control over his or her important outcomes. Such outcome dependency has been shown to increase attention to stereotype- or expectancy-inconsistent information and to result in more individuated impressions (Erber & Fiske, 1984; Fiske & Neuberg, 1990). Similarly, Neuberg (1989) documented how subjects given motivations for greater accuracy in their judgments (through experimental instructions) are more likely to overrule expectations and confirmatory hypothesis testing biases through a more complete gathering of individuating information. In several studies, Tetlock and his colleagues (e.g., Tetlock, 1985; Tetlock & Kim, 1987) showed how making subjects feel accountable for their impressions or judgments—"in that they believe they will have to defend and justify those judgments later—results in greater attention to situational constraints on the target’s behavior and, in general, more effortful decision making.

Finally, two recent studies showed that motivations can override the influence of passive priming effects on impression formation. Sedikides (1990) found that the saying is believing effect (Higgins & Rhees, 1978)—the tendency to shape one’s communication to fit the known beliefs or opinions of one’s audience, which then causes one’s judgments of the target to fall in line with those communications—overrode prior trait construct priming effects on subject’s impressions of the target’s ambiguous behaviors. Thompson et al. (1992) found that making subjects accountable for their judgments prior to reading about a target person even prevents subsequent priming effects on impressions.

Internally Generated Motivations

In the above studies, the source of the motivation to process effortfully resided in the situation, as manipulated by the experimental instructions. However, often the source of the motivation may be within the individual. D’Agostino and Fischer-Kiefer (1992) showed that subjects who were high in need for cognition (Cacioppo & Petty, 1982) are less likely to show the correspondence bias in attribution than other subjects, presumably because they chronically expend greater effort in mental processing and so are more likely to notice and use situational reasons for the target’s behavior. In Devine’s (1989) Experiment 3, subjects who valued not being prejudiced controlled the stereotypic content of their descriptions of
the average African-American, whereas prejudiced subjects did not (producing a more stereotypic description of that group).

Fiske and Von Hendy (1962) and Pittman and D'Agostino (1989) used experimental manipulations to activate motivations within subjects, which then determined how effortfully they formed an impression of a target person. Fiske and Von Hendy (1992) gave subjects who were either high or low self-monitors feedback about their person perception abilities (i.e., that they were good categorizers or individuators) as well as advice on the situational norms or appropriateness of categorizing or individuating people when forming impressions. The dispositional feedback determined whether low self-monitors (who presumably use their internal states, opinions, and abilities as behavioral guides) categorized or individuated the target, whereas the situational norm feedback determined whether high self-monitors (who are presumed to be more likely to use situational cues to guide appropriate behavior) categorized or individuated. Pittman and D'Agostino (1989) increased some subjects' control motivation by depriving them of control over their outcomes in an early part of the experiment. In a subsequent impression formation task, those subjects engaged in more effortful and careful processing of the information. This suggests that a more accurate prediction furnishes the perceiver with better predictive control over his or her environment.

Conclusions

Automated social cognitive processes categorize, evaluate, and impute the meanings of behavior and other social information, and this input is then ready for use by conscious and controlled judgment and decision processes, yet these judgments and decisions are not uncontrollable or predetermined by that automatic input. In the same way, the unintentional and uncontrollable nature of automatic analyses of the environment does not mean they are impossible to control or adjust for when one is aware of them, if one desires. Just as in strong perceptual illusions, one does not have to act in line with what is clearly (but inaccurately) apparent to one's senses when one knows better. The considerable body of research on motivational control over stereotypes and other judgmental biases has shown that, for the most part, the use of automatically supplied input in consciously produced judgmental output is not mandatory (see Bargh, 1988; Fiske, 1989; Jacoby & Kelley, 1990; Thompson et al., 1992).

THE AUTOMATICITY OF EVERYDAY LIFE: AN AGENDA FOR THE NEXT 10 YEARS

Automatic processes are not an unqualified blessing, nor are they an unqualified curse (Higgins & Bargh, 1992). Because of them, people stereotype others and often misunderstand the reasons for their own feelings and behavior. At the same time, automatic affective appraisal of the environment seems to be a ubiquitous and adaptive service (Bargh et al., 1992, 1993; Pratto & John, 1991). For example, Wilson and Schooler (1991) provided evidence that such immediate reactions may be more accurate (in terms of matching experts' opinions) than the preferences one comes up with after additional deliberation.

The automatization of routine thought processes frees one's limited attentional resources for nonroutine matters, and enables a reduction of the massive amount of stimulation and information bombarding one at any given moment into a more manageable subset of important objects, events, and appraisals. But with the increased efficiency of thought also comes a lack of awareness of engaging in that process, leading to a likelihood of misattributing the causes of one's feelings and a loosening of one's intentional grip over decisions and judgments.

Therefore, the recent wave of research demonstrating one's ultimate control over automatic input and judgmental processes is reassuring. In a sense, one is able both to delegate control to these automatized perceptual and judgmental mechanisms through frequent and consistent use of them in the past (Bargh & Gollwitzer, in press), and at the same time largely retain the final say over one's responses to the environment (Bargh, 1990; Logan & Cowan, 1984). Still, whether one exercises this ultimate control is another matter, as to do so one must be aware of the existence of the automatic influence, have the intention to effortlessly override it, and also sufficient attentional resources.

The research of the past 10 years has made it clear that the outcomes of social cognitive processes are very different, depending on whether one is aware of influences, whether one has specific intentions or goals within the situation, whether attentional resources are in ample or short supply, and whether one is motivated to take control over one's decisions and behavior. These are distinct and important dimensions on which social situations can vary. The past 10 years of research has been increasingly sensitive to the natural ecology of those situations with regard to a person's awareness, intentions, processing efficiency, and exercise of control within them. Should this trend continue, in the next decade even more will be discovered about the automaticity of everyday life.

ACKNOWLEDGMENTS

Preparation of this chapter was supported in part by Grant MH43265 from the National Institute of Mental Health. My thanks to Dan Gilbert, Roger Giner-Sorolla, Eliot Smith, Thom Srull, Erik Thompson, Jim Uleman, Bob Wyer, and Adam Zuckerman for their extensive comments on a previous version.

REFERENCES


1. AUTOMATICITY IN SOCIAL COGNITION


1. AUTOMATICITY IN SOCIAL COGNITION


