Journal of Personality and Social Psychology
2005, Vol. 88, No. 2, 245–262

0022-3514/05/$12.00 DOI: 10.1037/0022-3514.88.2.245

Violations of Implicit Theories and the Sense of Prediction and Control: Implications for Motivated Person Perception

Jason E. Plaks
University of Washington

Heidi Grant
Lehigh University

Carol S. Dweck
Stanford University

Beginning with the assumption that implicit theories of personality are crucial tools for understanding social behavior, the authors tested the hypothesis that perceivers would process person information that violated their predominant theory in a biased manner. Using an attentional probe paradigm (Experiment 1) and a recognition memory paradigm (Experiment 2), the authors presented *entity theorists* (who believe that human attributes are fixed) and *incremental theorists* (who believe that human attributes are malleable) with stereotype-relevant information about a target person that supported or violated their respective theory. Both groups of participants showed evidence of motivated, selective processing only with respect to theory-violating information. In Experiment 3, the authors found that after exposure to theory-violating information, participants felt greater anxiety and worked harder to reestablish their sense of prediction and control mastery. The authors discuss the epistemic functions of implicit theories of personality and the impact of violated assumptions.

What tools do people have at their disposal to interpret, explain, and predict human behavior? Recent research has shown that people often call on "implicit," "lay," or "naive" theories of personality (Hong, Levy, & Chiu, 2001; Morris, Ames, & Knowles, 2001). Two examples are the entity theory (the belief that human attributes are fixed and largely resistant to change) and the incremental theory (the belief that human attributes are malleable and cultivatable). A growing literature has demonstrated that these two theories precipitate distinct, contrasting patterns of social perception and explanation, one based on fixed traits as the primary vehicle for understanding behavior (the entity model) and one based on dynamic, psychological processes (the incremental model; for reviews, see Levy, Plaks, Hong, Chiu, & Dweck, 2001; Plaks, Levy, Dweck, & Stroessner, 2004).

The present research asked what occurs when people encounter information that violates their predominant theory. In the present studies, by providing participants with theory-violating information and observing the cognitive, affective, and motivational consequences, we examined the often implied but largely untested idea that implicit theories play a central role in maintaining people's sense of prediction and control competence.

Jason E. Plaks, Department of Psychology, University of Washington; Heidi Grant, Department of Psychology, Lehigh University; Carol S. Dweck, Department of Psychology, Stanford University.

Portions of this research were presented at the annual meeting of the Society for Personality and Social Psychology, February 2003, Los Angeles, California. We thank Tony Greenwald and Dan Molden for constructive feedback on earlier versions of this article.

Correspondence concerning this article should be addressed to Jason E. Plaks, Department of Psychology, University of Washington, Guthrie Hall 351525, Seattle, WA 98195-1525. E-mail: jep9@u.washington.edu

A second purpose of the present research was to illustrate that consistency with one's theory of personality is conceptually distinct from consistency with an expectancy or stereotype. That is, not all violations of expectancies—even strongly held ones—will trigger the same sense of control loss or the same defensive processing. This distinction, we argue, has important implications for basic stereotyping and person memory phenomena. In particular, it sheds light on when people will and will not be threatened by stereotype-disconfirming information.

A Motivational Approach to Implicit Theories

Inspired by cognitive psychology research on categorization (e.g., Kim & Ahn, 2002; Murphy & Medin, 1985), social psychological research has tended to describe implicit theories in primarily "cold" terms, as assumptions or heuristics that aid the classification of objects, people, and events (e.g., Morris et al., 2001; Skowronski, 2002; Wittenbrink, Hilton, & Gist, 1998). In this article, we suggest that besides serving this crucial cognitive function, implicit theories are vitally linked to "warm," epistemic motivations. We propose that core implicit theories such as the entity theory and the incremental theory play a key role in establishing and maintaining people's subjective sense of prediction and control competence. As such, people are heavily invested in believing that the theory they are using is correct. That is, rather than viewing their theory solely as a rule of thumb (that may be refined or discarded), people may adopt motivated processing distortions (e.g., selective attention, selective scrutiny) to protect the theory they are using from disconfirming evidence.

Although the assumption that people are motivated to maintain

a sense of prediction and control mastery implicitly underlies numerous strands of social psychological research, including the literatures on attribution (e.g., Burger & Hemans, 1988; Kelley, 1967), cognitive consistency (e.g., Swann, 1990), depression (e.g., Abramson, Seligman, & Teasdale, 1978; Gleicher & Weary, 1991), and stereotyping (e.g., Weary, Jacobson, Edwards, & Tobin, 2001), the assumption itself has received less systematic inquiry than might be expected (for exceptions, see Weary, Gleicher & Marsh, 1993; Pittman & D'Agostino, 1989). The present research addressed two critical, unanswered questions about prediction and control motivation and its implications for person cognition: (a) From which core beliefs about human nature might people's predictions arise, and (b) how do people respond when their predictions have been disconfirmed and their basic, working model of human personality has been violated?

We propose that although the entity and incremental theories are not the only sources of predictions about behavior, they are likely primary sources. As such, we hypothesized that (a) if predictions about human behavior often follow from one's predominant theory and (b) if people are motivated to believe that their theory is generally accurate, then (c) observing a person whose behavior violates these predictions should be a disorienting experience that triggers defensive processing and negative affect. Indeed, George Kelly (1955) drew an explicit connection between exposure to information that violates one's "personal construct" system and an increase in anxiety.[1] More recently, Weary and colleagues (e.g., Gleicher & Weary, 1991) demonstrated a link between a generalized feeling of causal uncertainty and depression, and Janoff-Bulman (1992) proposed that the psychological damage wrought by traumatic experiences may stem in large measure from the pain of "shattered assumptions." None of these studies, however, have explicitly or systematically investigated people's guiding principles of human personality and how the violation of such beliefs may lead people to feel less competent at predicting human behavior. Thus, the primary purpose of this research was to demonstrate that the entity and incremental theories not only belong to the class of beliefs that confer a sense of prediction and control but that people will adopt motivated processing distortions to protect them from disconfirmation.

It is crucial to note that we do not predict the experience of threatened prediction and control to follow automatically from the violation of *any* belief, expectancy, or stereotype (Olson, Roese, & Zanna, 1996). It is our contention, however, that implicit theories of personality cut to the very heart of what is meant by prediction and control; thus, their violation should be an epistemically disorienting experience. We also note that implicit theories are distinct from attitudes; whereas people may defend their attitudes because they are linked with deeply held values or ideologies (Eagly, Chen, Chaiken, & Shaw-Barnes, 1999), an implicit theory does not necessarily possess an evaluative aspect. We suggest, rather, that people are motivated to protect their implicit theories because they lend structure and predictability to a highly unpredictable social world.

Fixed Versus Malleable Theories Generate Distinct Predictions About Human Behavior

Consistent with work by psychologists and philosophers (e.g., Johnson, Gerner, Efran, & Overton, 1988; Pepper, 1942; Whitehead, 1929), Dweck and colleagues have proposed and found that most people possess basic theories about the fixedness or malleability of human personality (e.g., Dweck, 1999; Dweck & Leggett, 1988; Dweck, Chiu, & Hong, 1995). The entity theory's assumption that personal characteristics are fixed entities, despite a person's efforts or motivation to change them, is captured in the item from the Implicit Person Theories Measure, "Everyone is a certain kind of person, and there is not much that can be done to really change that" (Levy, Stroessner, & Dweck, 1998). The incremental theory's assumption that personal characteristics are dynamic and cultivatable with time and effort is reflected in the item, "Anyone can change even their most basic qualities."

Note that the entity and incremental theories are alternative lay perspectives on human nature, with neither necessarily reflecting the "correct" social reality. In addition, these theories can be both measured as chronic structures (using the Implicit Person Theories Measure; for validation information, see Levy et al., 1998) and situationally activated (using persuasive written materials) with essentially identical effects (Chiu, Hong, & Dweck, 1997, Experiment 5; McConnell, 2001, Experiment 2; Plaks, Stroessner, Dweck, & Sherman, 2001, Experiment 3). This suggests that both theories are intuitive, or available, to most people, although individuals may differ in which of the two theories is more chronically accessible (Levy et al., 2001).

The entity and incremental perspectives have been shown to influence cognition, affect, motivation, and behavior across a wide range of academic, social, and moral domains (for reviews, see Dweck, 1999; Dweck et al., 1995; Dweck & Leggett, 1988). For example, focusing on the domain of person perception, entity theorists, believing that personality traits are essentially fixed, tend to understand people and their behavior in terms of underlying traits (Chiu et al., 1997). This framework implies an expectation of high consistency in people's behavior over time and across situations; that is, if a person is "aggressive," he or she will behave more aggressively than average in the majority of relevant situations (Chiu et al., 1997; Erdley & Dweck, 1993; Molden, Plaks, & Dweck, 2004a). For incremental theorists, on the other hand, social understanding is not limited to diagnosing people's underlying fixed traits. To capture this more dynamic understanding of personality, incremental theorists assign relatively greater explanatory importance to mediating psychological and situational forces (Chiu, 1994; Hong, 1994; Molden, Plaks, & Dweck, 2004b). These assumptions, in turn, lead to the expectation that an individual's behavior can (and often does) vary significantly from situation to situation and over time (e.g., Chiu et al., 1997).

Building on these ideas, Plaks et al. (2001) used three established methods for assessing participants' attention to and encoding of social information (dichotic listening, attentional probe, recognition memory) and consistently found that entity theorists allocated less attention to stereotype-inconsistent information (e.g., a neo-Nazi skinhead who "helped an old lady across the street") than to consistent information (e.g., a skinhead who "trampled his

[1] As a clinician, Kelly's (1955) primary interest was in diagnosing and understanding each client's highly idiosyncratic system of personal constructs. We suggest, however, that certain basic ways of understanding human nature are shared by many individuals in a given culture and thus may be considered widely held "theories."

neighbor's flower garden")—a *congruency effect*. In contrast, incremental theorists paid more attention to stereotype-inconsistent information than to consistent information—an *incongruency effect*. Thus, Plaks et al.'s data demonstrated that rather than generally displaying a congruency effect (e.g., Bodenhausen, 1988; Stangor & Duan, 1991) or an incongruency effect (e.g., Bartholow et al., 2001; Hastie & Kumar, 1979; Macrae, Bodenhausen, Schloerscheidt, & Milne, 1999; J. W. Sherman, Lee, Bessenoff, & Frost, 1998), perceivers' preferences for consistent or inconsistent information depend in large measure on their a priori theories of personality.[2]

It should be noted that the processing preferences displayed by entity and incremental theorists were, as expected, found primarily under conditions of high cognitive load, consistent with recent findings and theoretical formulations (Macrae et al., 1999; J. W. Sherman et al., 1998; J. W. Sherman & Frost, 2000). According to the encoding flexibility model (J. W. Sherman et al., 1998), it is precisely when processing resources are scarce—forcing perceivers to decide which pieces of information are most deserving of attention—that such automatic preferences should emerge. (For a detailed treatment of this concept, see J. W. Sherman & Frost, 2000.) On the other hand, when perceivers have a fuller complement of resources at their disposal, thereby reducing the need to "make the tough choices" of what information to attend to, such preferences do not occur, and more evenhanded processing ensues. (However, such preferences may reemerge in extremely relaxed processing conditions, a phenomenon we demonstrate and explain in Experiment 2.)

Stereotype Violation Versus Theory Violation: A Motivational Side to Person Memory

One explanation for the difference between entity and incremental theorists' attention allocation involves the two theories' different assumptions about what kind of behavior (consistent vs. inconsistent) is more informative or *diagnostic* of an individual's personality (Skowronski, 2002). According to this diagnosticity-based explanation, because the entity theory carries the expectation that a person's behavior will be largely consistent, entity theorists consider trait-consistent behavior more informative than inconsistent behavior, which may be ascribed to random variation. Thus, consistent information is considered more worthy of one's finite processing resources. In contrast, because the incremental theory carries the expectation that a person's behavior can be quite variable, incremental theorists allocate more attention to inconsistent information, presumably because it can represent highly diagnostic information helpful for gaining a more complete portrait of a person (Skowronski, 2002).

We suggest that although such diagnosticity considerations play a key role in determining how elaborately perceivers will process consistent and inconsistent information, an additional, overlooked aspect may be perceivers' motivated attempts to decrease exposure to aversive, theory-violating information (e.g., Dijksterhuis, van Knippenberg, Kruglanski, & Schaper, 1996; Driscoll, Hamilton, & Sorrentino, 1991). In other words, does the presence of stereotype-inconsistent information violate entity theorists' theory and hence undermine their sense of prediction and control competence? If so, is there also a class of information that violates the incremental theory and hence undermines incremental theorists' sense of prediction and control?[3]

Consider the stereotype, well known to most college students, of the "math/science geek." Members of this group are presumed to possess very strong math and science skills and comparatively weak literary and artistic skills. We hypothesized that stereotype-inconsistent information (e.g., a math geek who gets a low score on the Math section of the Graduate Record Exam [GRE]) violates the entity theory because it contradicts the belief that defining traits, such as one's type of intelligence, are the stable, highly predictive building blocks of personality and behavior. In contrast, a math geek performing badly on a math exam does not necessarily violate the core assumption of the incremental theory (i.e., that people's behavior is dynamic and more susceptible to temporary, situational forces). Thus, entity theorists should be generally motivated to avoid or debunk counterstereotypic information, whereas incremental theorists should have no compelling reason to avoid or debunk such information and in fact should welcome it because of its potential usefulness.

To take these ideas one step further, it may be that in fact not all counterstereotypic behavior is equally threatening to entity theorists. Perhaps, to be truly threatening, the behavior must violate the presumed essential or defining traits of the target (e.g., "Brad [math geek] scored a 460 Math and 750 Verbal on a GRE exam"). However, behavior that is highly unexpected but not in violation of the target's core, defining trait (e.g., "Brad [math geek] eagerly renewed his subscription to *The New Yorker*") may be viewed as less threatening. Presumably, math intelligence is more central to the definition of math geeks than taste in magazines. (This intuition is verified in the pilot testing described below.) Thus, our hypothesis was that entity theorists would process counterstereotypic information that directly violated their theory more selectively than information that was highly counterstereotypic but did not violate their theory.

If so, this would have noteworthy implications for person memory research. Whereas typical studies in this literature have tended to operationalize the consistency of a target's behavior with respect to a focal trait or stereotype (e.g., an "intelligent" person who performs intelligent and unintelligent behaviors; Hastie & Kumar, 1979), our approach suggests that perceivers also define consistency with respect to their implicit theory of personality, in essence asking themselves, "Does this person's behavior support or violate my working model of personality?" Thus, entity theorists may not

[2] Studies have demonstrated that entity theorists' greater emphasis on traits and stereotypes is not explained by such person variables as higher right-wing authoritarianism (Altemeyer, 1988), simpler attributional complexity (Fletcher, Danilovics, Fernandez, Peterson, & Reeder, 1986), weaker intellectual ability (as indexed by grades and SAT scores), or higher general prejudice (Levy et al., 1998; Plaks et al., 2001, Experiment 3).

[3] A defense-motivational account is not intended to supplant possible nonmotivational accounts. There is substantial evidence that the processing of consistent and inconsistent information is associated with a number of purely cognitive biases that do not invoke any motivation beyond that of simply forming an impression (e.g., Johnston, Hawley, Plewe, Elliott, & DeWitt, 1990; Macrae, Bodenhausen, Schloerscheidt, & Milne, 1999; J. W. Sherman et al., 1998). We view the motivation to reach a particular conclusion as operating in tandem with such processes.

display biased processing toward all inconsistent behavior (and, as we describe below, incremental theorists may not display biased processing toward all consistent behavior). The notion that theory consistency may be different from stereotype consistency helps to define more clearly which types of behavior perceivers (as opposed to experimenters) view as inconsistent or undesired. This, in turn, leads to more precise predictions about who will attend more to which types of consistent and inconsistent behavior.

What Type of Information Violates the Incremental Theory?

Given that focusing on stereotype-consistent information helps to perpetuate stereotype belief and use (e.g., Bodenhausen, 1988), does it therefore follow that incremental theorists, by being more receptive to stereotype-inconsistent information, are by definition more accurate and open-minded person perceivers? According to the present approach, the answer to this question is "not necessarily." If it is truly the case that implicit theories are crucially involved in people's sense of prediction and control, then incremental theorists should exhibit similar defensive processes when their theory is violated.

How might the incremental theory be violated? With their belief in the dynamic and cultivatable nature of personality, incremental theorists may experience uneasiness on learning about a target person who is unable to change from a predetermined fate (i.e., someone who is "too" consistent). Consider, for example, a prototypical "math phobic" who is highly proficient in humanities subjects but largely helpless in mathematics and sciences. Imagine that this individual wholeheartedly undergoes an intensive, well-reputed program to improve his mathematics skills but in the end shows no improvement. If incremental theorists are also highly invested in believing that their theory is accurate, then such information about a person exhibiting rigid and predetermined personality—reflecting an inability to cultivate new skills or behavior—should be experienced as threatening and should, in turn, initiate defensive processing. (Note that incremental theorists do not believe people must change; instead, they are hypothesized to display biased processing toward information suggesting a person cannot change.)

Multiple Routes to Theory Protection

If theory protection is truly a fundamental motivation for both entity and incremental theorists, then there may be multiple routes to accomplish it, such that if one route (e.g., defensive inattention) is not feasible, a second (e.g., intensified scrutiny) may be used. Eagly and colleagues, in the context of the persuasion literature, have referred to this distinction as "passive" versus "active" defense (e.g., Eagly et al., 1999; Eagly, Kulesa, Chen, & Chaiken, 2000). As we describe in Experiment 2, a key predictor of whether such people will select active versus passive defense may be processing capacity. Whereas passive defense (e.g., selective attention and encoding) has been shown to be relatively resource independent (e.g., Bonanno, Davis, Singer, & Schwartz, 1991; MacLeod, Matthews, & Tata, 1986; Newman & McKinney, 2002), active defense (e.g., scrutinizing, debunking) may require plentiful resources (Ditto, Scepansky, Munro, Apanovitch, & Lockhart, 1998). Thus, when provided with sufficient processing capacity, theory-violating information may elicit more rather than less elaborate processing (i.e., approach rather than avoidance). In Experiment 1, we provided participants with highly versus moderately taxing processing conditions, expecting to find evidence of passive defense in the highly taxing conditions. In Experiment 2, we included three processing conditions: highly taxing, moderately taxing, and relaxed. We expected the highly taxing condition participants to blunt the impact of theory-violating information by using passive, avoidance strategies but the relaxed-condition participants to counter the same information with active, approach strategies.

Further Affective and Motivational Consequences

We hypothesized two further consequences of theory violation, one affective and one motivational. First, consistent with Kelly's (1955) analysis of "personal constructs," we expected an encounter with theory-violating information to lead to a rise in anxiety, because the individual's sense of prediction and control mastery has been threatened. Second, we hypothesized that the greater the anxiety participants experienced following theory violation, the more effort they would put into a subsequent task that enabled them to restore their sense of prediction and control. Previous research has indicated that people believe that using more effortful processing strategies can increase the likelihood of reaching subjectively "correct" judgments, thereby reasserting their mastery over the environment (e.g., Pittman & D'Agostino, 1989; Pittman & Pittman, 1980; see also Edwards & Weary, 1993; Walker & Sorrentino, 2000). On the basis of this research, in Experiment 3, we tested the hypothesis that participants would attempt to repair the anxious experience of control deprivation by redoubling efforts to assert their sense of control in another domain.

Overview of the Present Experiments

The present experiments addressed the following questions. (a) Who is more likely to be threatened by trait-inconsistent behavior (as opposed to other types of inconsistent behavior)? (Predicted answer: entity theorists.) (b) Who is more likely to be threatened by behavior implying excessively rigid trait consistency? (Predicted answer: incremental theorists.) (c) What strategies do people use to blunt the impact of theory-violating information? (Predicted answer: avoidance when processing resources are scarce, approach when processing resources are plentiful, and post hoc restoration of control when exposure to theory-violating information was persistent and unavoidable.)

To test these hypotheses, we conducted three experiments in which we presented participants with a series of behaviors performed by a target person that directly violated, confirmed, or was neutral with respect to their predominant implicit theory. We expected that both entity and incremental theorists would selectively avoid (Experiments 1–2) or scrutinize (Experiment 2) only those pieces of information that directly violated their respective theories. In Experiment 3, we tested for evidence of the proposed link between theory violation, increased anxiety, and the experience of control deprivation using a paradigm that measures participants' effort to repair a damaged sense of prediction and control (D'Agostino & Pittman, 1982). In sum, with these experiments, we intended to draw a detailed portrait of the experience of implicit

theory violation by (a) identifying who finds what type of information threatening and (b) highlighting major cognitive, affective, and motivational consequences.

Experiment 1

Previous researchers have used an attentional probe paradigm to assess how elaborately participants process different types of stimulus information (e.g., Hashtroudi, Mutter, Cole, & Green, 1984; MacLeod, Matthews, & Tata, 1986; Plaks et al., 2001, Experiment 1; J. W. Sherman, Conrey, & Groom, 2004; J. W. Sherman et al., 1998, Experiment 2). In Plaks et al. (2001, Experiment 1), participants read about behaviors performed by "Robert," a neo-Nazi skinhead (for half of the sample) or a priest (for the other half of the sample). (This method was adapted from J. W. Sherman et al., 1998, Experiment 2.) Ten kind, 10 unkind, and 10 neutral behaviors were presented in random order. As noted, the overall pattern was that under high cognitive load, entity theorists exhibited faster reaction times (RTs; taken to indicate less processing elaboration) to the stereotype-disconfirming behaviors, whereas incremental theorists exhibited slower RTs to those behaviors. However, because the behaviors used in the Plaks et al. studies (and, indeed, virtually all similar experiments in the person memory tradition) confounded two types of inconsistency (stereotype inconsistency and entity theory inconsistency), it was difficult to determine whether the motivation to protect one's theory played any part in entity theorists' congruency effect. Thus, in Experiment 1, we created a set of stimulus behaviors aimed at teasing apart this confound by isolating two dimensions of consistency and inconsistency.

Two Types of Stereotypic and Counterstereotypic Information: Trait Versus Associate

As depicted in Figure 1, counterstereotypic behavior that implies significant inconsistency along a core trait dimension (e.g., "Brad [math geek] scored a 460 Math, 750 Verbal on a GRE exam") violates the entity theory and thus is predicted to engage entity theorists' defense-motivational system. It is possible, however, for a person to behave counterstereotypically without casting doubt on his or her core traits. Such behavior, though highly inconsistent, falls along an associate dimension (e.g., "Brad [math geek] eagerly renewed his subscription to *The New Yorker*").

Associate behaviors may be thought of as relating to interests and tendencies strongly associated with (or strongly not associated with) members of a group without necessarily reflecting underlying, defining traits. For example, although in perceivers' minds it would be highly surprising for Brad the math geek to be an avid reader of *The New Yorker*, this fact should not speak to his characteristic type of intelligence as directly as his scores on the GRE.

We hypothesized that even when inconsistent–trait behaviors and inconsistent–associate behaviors are equated for perceived typicality, extremity, and diagnosticity (see Pilot Tests 1–3B), inconsistent–trait behaviors violate the entity theory more directly than inconsistent–associate behaviors. Thus, entity theorists should evince more biased processing toward inconsistent–trait behaviors. On the other hand, behavior suggesting a rigid predetermination of personality may be threatening to incremental theorists. Because such a person might be considered "too" consistent on the trait dimension, we predicted that incremental theorists would exhibit

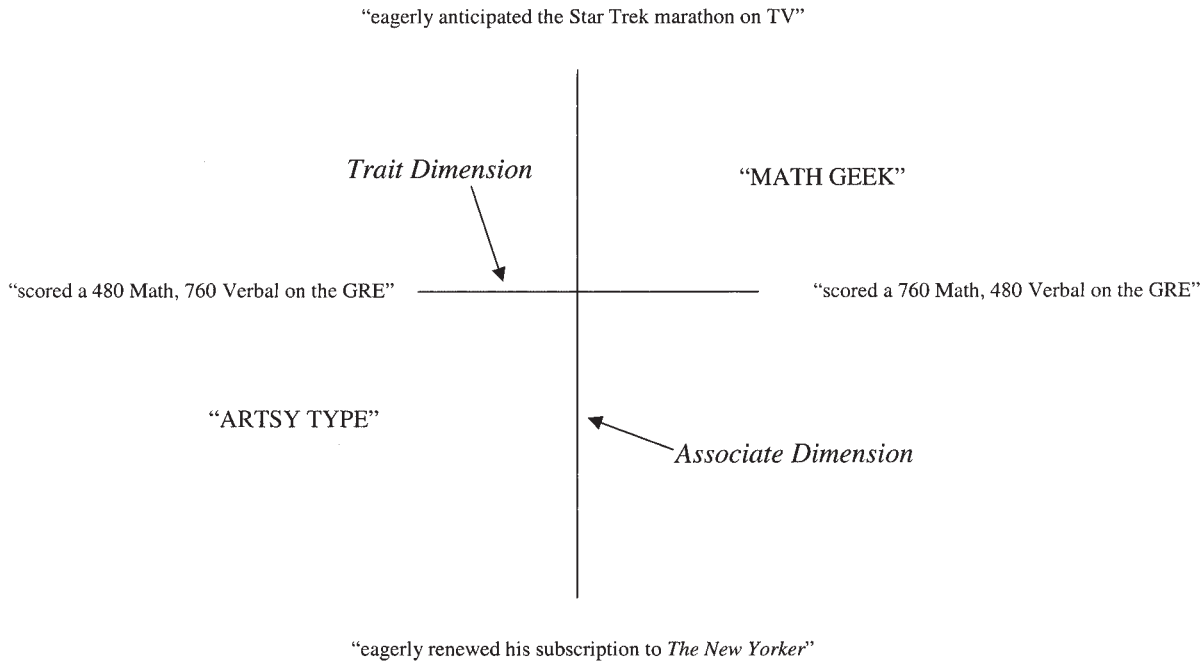

Figure 1. Two dimensions of stereotype-consistent and -inconsistent information about "math geeks" and "artsy types" (Experiment 1). GRE = Graduate Record Examination.

more defensive processing of consistent–trait behaviors than consistent–associate behaviors.

Pilot Studies

To conduct Experiment 1, it was necessary to conduct pilot studies to generate a set of consistent–trait, consistent–associate, inconsistent–trait, and inconsistent–associate behaviors. This allowed us to compare participants' attention with each type of behavior.

Pilot Study 1A. First, in Pilot Study 1A, we sought to verify that people indeed view type of intelligence as the primary dimension that distinguishes math and sciences people from arts and humanities people. We asked participants ($N = 29$) to rate the categories "math/sciences person" and "arts/humanities person" on a variety of trait dimensions, including "immoral/moral," "passive/active," and "math smart/humanities smart." As expected, these two categories differed only on the math smart/humanities smart dimension ($p < .01$).

Pilot Study 1B: The consistent–inconsistent dimension. In Pilot Study 1B, participants ($N = 45$) were asked to rate a long list of behaviors on the degree to which they were atypical (-5) or typical ($+5$) of "a math and sciences person" and a "humanities person." Examples included "scored a 460 on the Verbal GRE," "scored a 750 on the Math GRE," "got an A+ in European History," and "got an A+ in Linear Algebra." We selected the 30 behaviors that (a) had the highest difference scores between the math/sciences person and arts/humanities person subscales (i.e., highly consistent with one group and highly inconsistent with the other; mean difference score = 7.24, $SD = 2.48$) and (b) were rated equivalently by entity and incremental theorists. Thus, this pilot test yielded 15 behaviors that were math geek consistent–artsy person inconsistent and 15 behaviors that were artsy person consistent–math geek inconsistent.

Pilot Study 2: The trait–associate dimension. The next step was to divide the 30 stereotypic behaviors generated in Pilot Study 1B into trait dimension versus associate dimension behaviors. In Pilot Study 2A, participants ($N = 30$) were asked to rate these behaviors on the degree to which they "reflect someone's core, essential nature as a math and sciences person" or an "arts and humanities person" on a scale ranging from 1 (*completely reflective of a math/sciences person's core nature*) to 12 (*completely reflective of an arts/humanities person's core nature*) with a midpoint of 6 (*not reflective of a math/science person or an arts/humanities person's core nature*). The 12 behaviors that clustered toward the extremes were selected as trait behaviors (Ms > 9.17 and < 3.93, no entity-incremental differences). Examples included "scored a 760 on the Math GRE" and "got an A+ in European History." The 12 behaviors that clustered nearest the center of scale were labeled associate behaviors ($M = 6.26$). Examples included "went to a lecture given by a noted Shakespeare scholar" and "worked on a team that designs solar-powered cars."

In other words, the 12 associate behaviors were rated highly typical of arts/humanities people or math/sciences people but not necessarily reflective of their core nature. In contrast, the 12 trait behaviors were seen as more closely reflecting the person's underlying nature. For example, although it may be highly surprising for a math/sciences person to attend a lecture by a noted Shakespeare scholar, pilot subjects did not deem this a violation of his or her core nature. On the other hand, when such a person scores a 460 out of a possible 800 on the Math GRE (a low score in the eyes of our undergraduate population), this would be a violation of his or her presumed nature as a math geek.

In Pilot Study 2B, we compared whether the 24 newly designated trait and associate behaviors differed in degree of typicality. Each of the trait behaviors did not differ significantly from any of the associate behaviors in typicality (ts < 1.88, *ns*), ruling out differences in perceived typicality between the trait behaviors and the associate behaviors as a potential confound. Additionally, there were no differences between the ratings of entity and incremental pilot subjects (all ts < 1.30, ps $> .20$).

Pilot Study 3A: Behavior extremity. It was crucial to establish that the trait behaviors were not simply more "extreme, emphatic, or demonstrative" than the associate behaviors. A sample of participants ($N = 30$) rated the behaviors generated by Pilot Studies 1–2 for their extremity. None of the 12 trait behaviors were rated more extremely than any of the 12 associate behaviors (ts < 1.78, ps $> .14$). Moreover, as in each of the previous pilot studies, there were no significant differences between entity and incremental theorists on these behavior ratings (all ts < 0.99, ps $> .35$).

Pilot Study 3B: Behavior diagnosticity. According to Skowronski and Carlston (1987; Skowronski, 2002), *behavior diagnosticity* refers to a behavior's "ability to predict membership in one of two alternative trait categories" (Skowronski, 2002, p. 138). If the so-called trait and associate behaviors also differed in diagnosticity, this could present an alternative account for any observed differences. Thus, we closely followed Skowronski's (2002) method for establishing behavior diagnosticity by presenting participants ($N = 50$) with the behaviors generated in Pilot Studies 1–3A and asking them to rate "how useful each behavior is for deciding whether it was performed by a math and sciences person or an arts and humanities person." As in Skowronski's study, additional instructions asked participants to assume that the person performing the behavior had performed other behaviors, some of which may have been inconsistent with the one they had just read. (These instructions were intended to steer participants away from relying on the representative heuristic; see Skowronski, 2002.) Results indicated that the 12 trait dimension behaviors were not rated as more diagnostic than the 12 associate dimension behaviors (all ts < 1.01, all ps $> .25$; no entity-incremental differences). Thus, any differences in participants' attention to trait or associate behaviors in the experimental session could not be due to a priori differences in how diagnostic of math/sciences people and arts/humanities people the behaviors were perceived to be. The final set of 24 behaviors generated by the pilot testing (12 trait, 12 associate) was combined with 6 neutral–irrelevant behaviors (e.g., "went to the laundromat to do his laundry") to form a total stimulus set of 30 behaviors.

Experimental Session: Method

Participants. One hundred sixteen University of Washington undergraduates (84 women, 32 men) gave informed consent and participated for extra course credit.

Procedure. Participants completed the eight-item Implicit Person Theories Measure as part of a large questionnaire battery at the beginning of the term, several weeks prior to the pilot testing or the experimental session. Consistent with Kelly's (1955) direct approach to examining people's underlying theories of their social world, the Implicit Person

Theories Measure directly assesses participants' theories about the fixedness or malleability of human characteristics. (Though these theories have been termed "implicit" because they are usually poorly articulated, it is presumed that people are able to agree or disagree with the simple, straightforward items on our measure.) There are several forms of this measure corresponding to particular domains (e.g., intelligence, morality; see Dweck et al., 1995; Levy et al., 1998). In this experiment, the "intelligence" form of the measure was used because stereotypes of math geeks and artsy types pertain to different types of intelligence (see Dweck et al., 1995; Levy et al., 1998). The measure asks participants to rate their degree of agreement–disagreement with items such as "People can do things differently, but the important parts of their intelligence cannot really be changed" and "People can substantially change their intelligence" (reverse scored). Further discussion of issues regarding the reliability and validity of different versions of this measure can be found elsewhere (Levy et al., 1998).

In the experimental session, participants read a booklet containing a vignette about "Brad," an incoming college freshman. Brad was described as either a "math/sciences person" or an "arts/humanities person." The text provided participants with several details that supported this claim. For example, Brad's SAT scores (750 Math, 460 Verbal for the math and sciences target; 750 Verbal, 460 Math for the arts and humanities target), grades, hobbies, and academic interests were manipulated to convey either a person who was strong in math and sciences but weak in arts and humanities or a person who was strong in arts and humanities but weak in math and sciences.

Next, the vignette stated that "Brad's university requires all incoming freshmen with subpar backgrounds in humanities [mathematics] to undergo an intensive, year-long course in English language and expository writing [calculus and statistics]." Participants read that "Brad looks forward to the course and plans to try his best." Thus, this vignette created a scenario in which the target was afforded ample opportunity to develop from his prior academic disposition and cultivate new (counterstereotypic) skills.

Participants were then asked to imagine that currently it was a year later and Brad had completed the course. They were asked to rate "How good is Brad now at arts, humanities, and social sciences-related subjects?" and "How good is Brad now at math and hard sciences-related subjects" on scales ranging from 1 (*very, very poor*) to 7 (*very, very good*). These items served to verify whether entity and incremental theorists reached different expectancies for the target on the basis of the same background information. As predicted, when Brad was a math and sciences person, entity theorists rated him lower on the humanities scale ($M = 3.72$) than did incremental theorists ($M = 5.04$), and when Brad was a humanities person, entity theorists rated him lower on the math and sciences scale ($M = 3.36$) than did incremental theorists ($M = 4.49$; both ts > 3.56, both ps $< .05$). (This replicates a previous finding; Plaks, 2001, Experiment 3.) In other words, entity theorists were less convinced that the remedial course was successful in improving the target's ability.

Next, participants performed a computerized attentional probe task. Instructions stated that participants would read a representative sample of behaviors performed by Brad during the course of a recent week. Then, the computer presented a randomized sequence of the behaviors generated by Pilot Tests 1–3. There were six instances of five categories of behavior: consistent–trait (e.g., "Brad [artsy] got a 760 Verbal, 480 Math on the GRE"), consistent–associate (e.g., "Brad [artsy] eagerly renewed his subscription to *The New Yorker*"), inconsistent–trait (e.g., "Brad [artsy] got a C− in an easy European History class"), inconsistent–associate (e.g., "Brad [artsy] joined the university's team that designs and builds solar powered cars"), and neutral (e.g., "Brad [artsy] went to the laundromat to do his laundry"). Sentences were constructed to be all nearly the same length.

As the behaviors were being presented (one at a time for 3.5 s each), the computer emitted a tone 2.0 s into the presentation of some of the behaviors. (These values were generated through pilot testing described by J. W. Sherman et al., 1998, Experiment 2.) A total of 10 tones were emitted, twice for each of the five types of behaviors. The computer randomized which behaviors within each type received a tone as well as the overall order of presentation of the behaviors. Participants were instructed to press the space bar as quickly as possible after hearing a tone. The computer measured response latency. According to the logic of this paradigm, the less the participant is engaged with the sentence currently on the screen, the faster the RT will be, because it is easier to disengage from a stimulus that is not fully absorbing (Hashtroudi et al., 1984; Plaks et al., 2001, Experiment 1; J. W. Sherman et al., 1998).

As in previous experiments (e.g., Plaks et al., 2001, Experiment 1; J. W. Sherman et al., 1998, Experiment 2; J. W. Sherman et al., 2004), roughly half of the participants were randomly assigned to the high cognitive load condition. These participants were instructed to count backward from 938 by sevens out loud as they were viewing the sentences on the computer screen. This was done because numerous studies (e.g., Macrae et al., 1999; Plaks et al., 2001; J. W. Sherman et al., 1998, 2004) have found that encoding preferences for stereotype-consistent and -inconsistent information are particularly acute under conditions of high cognitive load—an effect predicted by J. W. Sherman et al.'s (1998, 2004) encoding flexibility model. After completing the attentional probe task, participants were fully debriefed.[4]

Hypotheses. Information that counters a target's defining characteristic (e.g., a "math person" who suddenly cannot do math) should violate the entity theory. In contrast, information that counters a target's ability to learn and develop his weakness (e.g., a "math person" who still cannot read and write well, even after trying hard in a rigorous remedial writing course) should violate the incremental theory. Thus, we predicted that under high cognitive load, (a) entity theorists would exhibit faster RTs (indicating less processing elaboration) to inconsistent–trait behaviors than to consistent–trait behaviors, but this preference would diminish for behaviors on the associate dimension; (b) incremental theorists would exhibit faster RTs to consistent–trait behaviors than to inconsistent–trait behaviors, but this difference would diminish for behaviors on the associate dimension.

Results and Discussion

Responses to the Implicit Person Theories Measure. Participants' responses to the Implicit Person Theories Measure items were highly reliable (Cronbach's $\alpha = .93$). Accordingly, responses to the eight items were averaged, after reverse scoring where appropriate, to create an implicit person theory index for each participant. As in previous research (e.g., Plaks et al., 2001), participants with a mean theory score of 3.0 or below (indicating overall agreement) were classified as entity theorists ($n = 47$), and participants with a mean score of 4.0 and above (indicating overall disagreement) were classified as incremental theorists ($n = 47$). Participants with mean theory scores that fell between 3.0 and 4.0 were unclassified ($n = 19$) and were excluded from the analyses.

[4] In addition, immediately prior to and immediately after reading the behaviors, participants completed a short questionnaire that assessed their current affective level. Although it would have been ideal to use these affect data to demonstrate that the arousal of anxiety mediated the relationship between information type and RT, such analyses were not appropriate with the present design because affective states such as arousal often linger for longer than the 3.5 s each behavior was on the screen (Zillman & Zillman, 1996), violating the assumption of causal transience necessary for mediational analyses in a within-subjects design (Judd, Kenny, & McClelland, 2001, p. 117). However, see Experiment 3 (which used a between-subjects design).

Table 1
Reaction Times (in Milliseconds) to Tone as a Function of Type of Information and Participants' Implicit Theory, Experiment 1 (All Values Rounded to Nearest Whole Number)

	Type of information				
	Trait dimension		Associate dimension		
Theory	Consistent	Inconsistent	Consistent	Inconsistent	Neutral
	High cognitive load				
Entity theorists					
M	853	659	866	810	792
SD	84	110	123	120	170
Incremental theorists					
M	738	801	839	795	859
SD	96	82	111	78	222
	Low cognitive load				
Entity theorists					
M	517	512	504	489	497
SD	129	118	99	105	108
Incremental theorists					
M	518	497	494	484	513
SD	174	159	131	146	112

Response latency analyses. After RTs greater than 2 standard deviations from the mean were removed (2.7%), mean RTs to the tones during the presentation of each type of information were computed for each participant. To test the relationship between participants' implicit theory and their RTs to each type of information, we conducted a 2 (theory: entity vs. incremental) × 2 (cognitive load: high vs. low) × 2 (consistency: consistent vs. inconsistent) × 2 (dimension: trait vs. associate) analysis of variance (ANOVA) with repeated measures on the last two factors. This analysis revealed main effects for cognitive load, $F(1, 87) = 191.62$, $p < .01$ (indicating the effectiveness of the load manipulation); consistency, $F(1, 87) = 11.01$, $p < .01$ (indicating that participants did not allocate equal attention to consistent and inconsistent information); and dimension, $F(1, 87) = 17.41$, $p < .01$ (indicating that participants did not allocate equal attention to trait and associate information). Testing our main hypothesis, the analysis found the predicted interactions for Theory × Consistency × Dimension, $F(1, 87) = 9.98$, $p = .002$ (indicating that entity and incremental theorists differed in their patterns of attention to information along the two dimensions), and Theory × Consistency × Dimension × Cognitive Load, $F(1, 87) = 14.14$, $p < .01$ (indicating that the three-way interaction varied by level of cognitive load).[5]

Recall our prediction that entity theorists under high cognitive load would display a stronger congruency effect (faster RTs to inconsistent information compared with consistent information) on the trait dimension than on the associate dimension, but incremental theorists (in the high cognitive load condition) would display a stronger incongruency effect on the trait dimension than on the associate dimension. Indeed, as depicted in Table 1, both entity theorists and incremental theorists displayed a significant (or near-significant) Consistency × Dimension × Cognitive Load interaction: entity theorists, $F(1, 90) = 7.05$, $p < .01$; incremental theorists, $F(1, 90) = 2.99$, $p < .09$. However, as predicted, the nature of these interactions differed. The Consistency × Dimension interaction for high cognitive load entity theorists, $F(1, 90) = 17.41$, $p < .01$, included a stronger congruency effect on the trait dimension, $F(1, 90) = 61.13$, $p < .01$, than on the associate dimension, $F(1, 90) = 10.75$, $p < .01$. In contrast, high cognitive load incremental theorists' Consistency × Dimension interaction, $F(1, 90) = 9.30$, $p < .01$, included a strong incongruency effect on the trait dimension $F(1, 90) = 5.67$, $p < .02$, but, if anything, a reversal on the associate dimension, $F(1, 90) = 3.71$, $p < .06$. Also as predicted, there were no significant effects for either entity theorists or incremental theorists in the low load condition (all Fs < 1.5).

Approach or avoidance? To determine whether participants used an approach or avoidance strategy toward theory-violating information, we compared high cognitive load participants' RTs to theory-violating and theory-confirming behaviors with their RTs to neutral behaviors. Entity theorists' RTs were significantly faster during inconsistent–trait (i.e., theory-violating) behaviors than during neutral (irrelevant) behaviors, $F(1, 90) = 16.32$, $p < .01$, but their RTs to consistent–trait (theory-confirming) behaviors were not significantly different from neutral information, $F(1, 90) = 2.45$, *ns*. (This avoidance tendency replicates a previous finding; Plaks, Grant, & Dweck, 2004.) Incremental theorists showed a parallel pattern, with faster RTs to consistent–trait information

[5] This analysis also yielded significant but less theoretically relevant interactions for Theory × Dimension, $F(1, 87) = 17.41$, $p < .01$, and Cognitive Load × Dimension, $F(1, 87) = 7.25$, $p < .01$. In addition, an analysis with type of expectancy (math person vs. humanities person) included as a factor contributed negligibly to the overall variance and is not discussed further. The three- and four-way interactions replicate a similar, unpublished study that used "skinhead" or "priest" as the target categories and had the target engage in varieties of moral and immoral behavior (Plaks, Grant, & Dweck, 2004).

than to irrelevant information, $F(1, 90) = 11.56$, $p < .01$, but equivalent RTs to inconsistent–trait information and irrelevant information, $F(1, 90) = 2.53$, *ns*. Thus, both groups appeared to pursue the strategy of avoiding theory-violating information rather than approaching theory-confirming information.

In sum, under high cognitive load, both entity and incremental theorists exhibited faster RTs only to behaviors that violated their respective theories. This can be taken as evidence that the phenomenon of selective attention in the name of theory protection applies to both entity theorists and incremental theorists and may well be a generally pervasive phenomenon.

As noted, the finding that perceivers distinguish between different varieties of consistent and inconsistent behavior has important implications for research on person memory (e.g., Hastie & Kumar, 1979; Srull, 1981; Stangor & McMillan, 1992). This literature has focused on how a behavior's consistency or inconsistency with a stereotype can influence the elaborateness with which it is processed. The present research suggests that people not only consider a behavior's consistency with a stereotype—they also consider its consistency with their theory of personality. Indeed, in Experiment 1, entity theorists avoided counterstereotypic behavior that also violated their theory more than behavior that was merely counterstereotypic, and incremental theorists avoided stereotypic behavior that also violated their theory more than behavior that was merely stereotypic.

Yet is selective exposure the only method perceivers have at their disposal to handle theory-violating information? It is likely that the processing picture is more complicated. If theory protection is truly a fundamental motivation, then people may possess multiple routes to accomplish it, such that if one route is inoperable, a second may be used. In many cases, undesired information may actually elicit more elaborate processing (e.g., Ditto et al., 1998; Eagly et al., 2000). In Experiment 2, we sought to identify one condition when this might occur.

Experiment 2

Avoidance of theory-violating information may at times be impractical or undesirable. Consider, for example, when information about a target person is presented in a manner that is extremely emphatic, persistent, or simple to process. In such a case, it may be difficult for people to avoid devoting significant resources to processing each piece of information. This does not mean, however, that biased processing cannot occur. When inescapably confronted with theory-violating information, perceivers may direct additional scrutiny toward the offending information as they attempt to fully understand it, cope with it, or perhaps debunk it (Chaiken, Giner-Sorolla, & Chen, 1996; Ditto et al., 1998; Förster, Higgins, & Strack, 2000; Hastie, 1984). This distinction between "passive" and "active" defense has been discussed in the persuasion literature (e.g., Eagly et al., 1999), although little contemporary research has attempted to isolate variables predicting when people will pursue one mode over the other (Baumeister, Dale, & Sommer, 1998; Chaiken et al., 1996; see also Frey, 1986).

There is reason to believe that level of cognitive load is one important variable that predicts whether people will pursue an active versus passive mode of motivated processing. Whereas passive processes such as selective attention are likely to be comparatively resource-independent (e.g., MacLeod, Matthews, & Tata, 1986; Newman & McKinney, 2002), active processes such as discounting or debunking are likely to require ample processing resources (e.g., Ditto et al., 1998; Förster et al., 2000). In Experiment 2, we investigated the possible role of processing resources by including a condition that provided participants with significantly more time to digest each of the presented behaviors than in even the low cognitive load condition of Experiment 1 (8.0 s vs. 3.5 s per sentence). That is, we introduced a new condition of cognitive load in this experiment: "no" load. Whereas in the low load condition of Experiment 1, participants were faced with the somewhat challenging task of reading each sentence in 3.5 s while monitoring for an auditory tone, in Experiment 2's no load condition, there was more than twice as much time to read each sentence and no auditory monitoring task. These comparatively relaxed processing conditions mimic those found in previous studies in which participants exhibited more (rather than less) elaborate processing of undesired information (e.g., Ditto et al., 1998; Förster et al., 2000). Moreover, Experiment 2 participants were able to turn back and revisit previous sentences (unlike in Experiment 1). We expected this method of presentation to have a dual effect: (a) It would make the option of filtering out a given sentence less likely because of each sentence's persistent duration, and (b) it would allow participants the opportunity to read carefully, and perhaps even reread, any particular sentence.

Is there an empirical precedent for this hypothesis of increased, rather than decreased, cognitive engagement with motivationally aversive stereotype-relevant information? In one pertinent study, Förster et al. (2000) presented participants with a list of stereotypic and counterstereotypic target behaviors at a rate of 6.67 s per sentence (with no concurrent task and no cognitive load) and then assessed participants' recognition sensitivity for each type of sentence after a delay (recognition sensitivity measures are widely assumed to tap the degree of thoroughness with which information was initially encoded; e.g., Eagly et al., 2000; J. W. Sherman & Frost, 2000; Srull, 1981). Förster et al. found that participants for whom counterstereotypic information was predicted to be especially threatening (according to their level of prejudice or their regulatory focus) exhibited greater recognition for counterstereotypic information. In other words, these participants did not screen out undesired information; rather, they encoded such information more elaborately. According to Förster et al., threatened participants gathered more detail to truly understand the threat and its context.[6] (For a more general treatment of this idea, see Ditto et al., 1998.)

When circumstances work against selective attention, might entity and incremental theorists attempt to protect their theory through extra elaboration? In Experiment 2 we tested this idea by using the same stimulus set as in Experiment 1 but included three cognitive load conditions: high load, low load, and no load (for a discussion of the merits of including more than two levels of cognitive load, see Spears & Haslam, 1997). In the no load

[6] Using a similar paradigm, J. W. Sherman and Frost (2000) did not find any meaningful effects in a low cognitive load condition that resembled Förster et al.'s (2000; 6 s per presented sentence). J. W. Sherman and Frost did not, however, measure any motivational or individual difference variables. Thus, implicit theory effects could in fact have been present but undetected in J. W. Sherman and Frost's low cognitive load condition data.

condition, we significantly relaxed the processing demands compared with the low load condition. If Experiment 2 participants showed evidence of flexibly adopting an alternate means of theory protection (increased rather than decreased encoding elaboration), this would suggest that the same underlying motivational–epistemic state may precipitate different encoding strategies depending on processing conditions. Such a finding would lend important detail to our understanding of the subtle mechanisms behind people's efforts to protect their implicit theories.

Another purpose for using a recognition memory paradigm in Experiment 2 to assess participants' degree of encoding elaboration was to verify that the findings in Experiment 1 were not limited to or a function of the attentional probe method.

Method

Participants. A total of 98 University of Washington students (61 women, 37 men) gave informed consent. The stimuli were presented to participants in paper-and-pen format.

Procedure. Participants completed the Implicit Person Theories Measure several weeks prior to the experimental session as part of a multiquestionnaire battery session. In the experimental session, participants read background materials about "Brad" (a student who, as in Experiment 1, was described as a "math/sciences person" or an "arts/humanities person" who took a year-long intensive course designed to improve his academic weaknesses). Next, participants read about the target's behaviors during a typical week. The 30 behaviors performed by the target were written on seven pages (four or five per page). At the appointed time, the experimenter instructed participants to begin reading the sentences carefully. The experimenter used a stopwatch and verbal instructions to measure and regulate the amount of time participants had to read the target behaviors.

Cognitive load manipulation. Participants were randomly assigned to one of three cognitive load conditions. In the high load condition, participants were given 2 min to read the sentences (i.e., 4 s per sentence). In addition, participants in this condition were required to memorize an eight-digit number and report the number back after reading the sentences. In the low load condition, participants were given 2 min to read the sentences but were not given an eight-digit number to memorize. In the no load condition, participants were given 4 min to read the sentences (8 s per sentence) and no number to memorize. (As noted, the parameters of this no cognitive load situation were considerably less demanding than even the low cognitive load condition in the present study as well as the low load condition of Experiment 1, in which each sentence appeared on the screen for only 3.5 s, and participants could not return to reread a sentence once it had passed.)

After a 10-min delay, participants were presented with a randomized list of 20 of the 30 behaviors plus 20 foil behaviors in random order. Participants' task was to indicate which behaviors they had seen before and not seen before by marking the appropriate columns. After completion of this task, participants were fully debriefed.

Results and Discussion

Responses to the Implicit Person Theories Measure. Participants' responses to the Implicit Person Theories Measure (intelligence version) items were reliable (Cronbach's $\alpha = .90$). As in Experiment 1, participants with a mean theory score of 3.0 or below (indicating overall agreement) were classified as entity theorists ($n = 34$), and participants with mean scores of 4.0 and above (indicating overall disagreement) were classified as incremental theorists ($n = 32$), with the remaining unclassifiable participants excluded from the analyses.

Recognition sensitivity data. For each participant, we calculated recognition sensitivity (the ability to discriminate hits from false alarms) for each of the five types of information, $d' = z(\text{hits}) - z(\text{false alarms})$ (Macmillan & Creelman, 1991). The data of 4 participants who obtained 100% accuracy (1 entity theorist, 2 incremental theorists, 1 unclassified) were excluded from the analysis because d' cannot be calculated from such data without additional assumptions and adjustments (Macmillan & Creelman, 1991). The remaining d' scores were entered into a 2 (theory: entity vs. incremental) × 3 (load: high, low, no) × 2 (consistency: consistent vs. inconsistent) × 2 (dimension: trait vs. associate) ANOVA with repeated measures on the last two factors. This analysis revealed an interaction for Theory × Dimension, $F(1, 60) = 16.10$, $p < .03$, indicating that entity and incremental theorists differed in their pattern of attention to trait versus associate information, and for Theory × Dimension × Load, $F(2, 60) = 14.84$, $p < .01$, indicating that this pattern differed across the three load conditions. These relationships were moderated by the predicted Theory × Load × Consistency × Dimension interaction, $F(2, 60) = 6.38$, $p < .01$. (As in Experiment 2, an analysis with type of target [math person vs. humanities person] as an additional factor contributed negligibly to the overall variance.)

Recall our prediction that as in Experiment 1, entity theorists under high cognitive load would display a stronger congruency effect (less elaborate encoding of inconsistent information than of consistent information) on the trait dimension than on the associate dimension but no such interaction under low cognitive load. Recall also the novel prediction of Experiment 2: that entity theorists' Consistency × Dimension interaction would reverse under no cognitive load (a stronger incongruency effect on the trait dimension than on the associate dimension). Indeed, this was the case. Replicating Experiment 1, entity theorists displayed a Load × Consistency × Dimension interaction under high cognitive load, $F(2, 60) = 7.56$, $p < .01$, as depicted in the top section of Table 2. Although the Consistency × Dimension interaction did not reach significance, $F(1, 60) < 2.0$, *ns*, planned comparisons revealed that as predicted, entity theorists displayed a stronger congruency effect on the trait dimension, $F(1, 60) = 4.16$, $p < .05$, than on the associate dimension, $F(1, 60) = 0.11$, *ns.* As depicted in the bottom section of Table 2, under no cognitive load, the Consistency × Dimension interaction exhibited by entity theorists, $F(1, 60) = 6.54$, $p < .02$, was made up of a stronger incongruency effect on the trait dimension, $F(1, 60) = 56.41$, $p < .01$, than on the associate dimension, $F(1, 60) = 6.69$, $p < .02$. This greater recognition accuracy suggests that entity theorists in the no load condition encoded inconsistent–trait information, which violated their assumption of fixedness, more elaborately than inconsistent–associate information, which did not violate the fixedness assumption.

Incremental theorists likewise displayed a near-significant Load × Consistency × Dimension interaction, $F(2, 60) = 2.49$, $p = .10$. Unlike entity theorists, however, incremental theorists' Consistency × Dimension interaction under high cognitive load, $F(1, 60) = 3.72$, $p < .06$, was made up of a stronger incongruency effect on the trait dimension, $F(1, 60) = 5.83$, $p < .02$, that was eliminated on the associate dimension, $F(1, 60) = 0.08$, *ns* (top section of Table 2). In line with the novel prediction for Experiment 2, under no cognitive load, the direction of incremental theorists' near-significant Consistency × Dimension interaction,

Table 2
Recognition Sensitivity (d′) as a Function of Cognitive Load, Type of Information, and Participants' Implicit Theory, Experiment 2

	Type of information			
	Trait dimension		Associate dimension	
Theory	Consistent	Inconsistent	Consistent	Inconsistent
High cognitive load				
Entity theorists				
M	4.64	3.09	4.33	4.02
SD	1.62	2.52	2.61	2.54
Incremental theorists				
M	2.21	4.86	5.30	4.85
SD	3.44	1.65	1.51	1.64
Low cognitive load				
Entity theorists				
M	4.14	5.29	3.53	5.74
SD	2.43	1.51	2.13	1.17
Incremental theorists				
M	4.49	3.93	3.37	4.21
SD	1.61	1.99	2.57	2.50
No cognitive load				
Entity theorists				
M	1.64	5.99	3.45	5.27
SD	2.22	0.75	2.65	1.82
Incremental theorists				
M	4.86	3.31	3.97	4.19
SD	1.59	1.90	2.82	1.96

$F(1, 60) = 2.72$, $p < .10$, reversed as they displayed a stronger congruency effect on the trait dimension, $F(1, 60) = 8.56$, $p < .01$, than on the associate dimension, $F(1, 60) = 0.16$, *ns* (bottom section of Table 2).[7]

Finally, as shown in the middle section of Table 2, although the patterns of means for participants in the low load condition generally resembled those in the no load condition in their direction, the key Consistency × Dimension interaction was not significant for either entity theorists, $F(1, 60) = 0.77$, *ns*, or incremental theorists, $F(1, 60) = 1.35$, *ns.* This pattern replicates Experiments 1 (as well as Plaks et al., 2001, Experiments 1–2; J. W. Sherman, Conrey, & Groom, 2004; J. W. Sherman et al., 1998), in which moderately relaxed processing conditions produced less biased processing of stereotype-relevant information. One potential explanation for this pattern is that in such "in-between" processing conditions, perceivers engage to some extent in both strategies (defensive inattention and defensive scrutiny), and these opposing tendencies cancel each other out. Another possible contributor is that participants in moderately taxing conditions (compared with highly taxing conditions) are better positioned to follow the stated or implied instructions to process each behavior relatively even-handedly. An interesting avenue for future research would be to tease apart the relative contribution of these different processes.

To summarize Experiment 2, changing from an attention allocation paradigm (Experiment 1) to a recognition memory paradigm (Experiment 2) yielded a similar pattern of effects, suggesting that both measures tapped similar underlying processes. Experiment 2 added to Experiment 1 by randomly varying between three processing conditions within the same study. This tri-level cognitive load manipulation revealed that load is a good predictor of whether people will adopt a more passive strategy or a more active, aggressive strategy of selective processing. Participants displayed poorer recognition accuracy for theory-violating behaviors (presumably because of decreased scrutiny) when processing resources were depleted but greater recognition accuracy to the same behaviors (presumably because of increased scrutiny) when processing resources were plentiful. Taken together, Experiments 1–2 provide evidence that people can flexibly use both avoidance and approach strategies to deal with theory-violating information.

In Experiment 3, we investigated a third strategy for blunting the impact of exposure to theory-violating information: post hoc efforts to restore the sense of prediction and control. More specifically, we tested the proposed relationship between exposure to theory-violating information, increased anxiety, and a threatened sense of prediction and control.

Experiment 3

Previous research on prediction and control motivation has found that when people lose confidence in their ability to predict and control events (whether because of a chronic sense of causal uncertainty or situational manipulations), they call on more methodical information-gathering strategies aimed at reestablishing a sense of prediction and control mastery (e.g., Pittman & D'Agostino, 1989; Pittman & Pittman, 1980; Weary, Jacobson, Edwards, & Tobin, 2001). Presumably, such a strategy is one of the compensatory processes people use to restore certainty to their self-identity after it has been cast into doubt (e.g., McGregor & Marigold, 2003). If it is the case, as we propose, that implicit theories play a crucial role in people's sense of prediction and control, then when people are unavoidably exposed to information that violates their theory, they should dedicate themselves to restoring their sense of prediction and control mastery.

In addition, on the basis of previous research illustrating the relationship between receiving undesired information and anxiety (e.g., Elliot & Devine, 1994; Förster et al., 2000), we predicted that exposure to theory-violating information would trigger an increase in anxiety for both entity and incremental theorists. We further hypothesized that this experience of anxiety would mediate participants' subsequent efforts to restore control.

To investigate this possibility, we adapted an established paradigm that assesses control-deprived participants' efforts to regain control (D'Agostino & Pittman, 1982; cf. Alloy & Abramson, 1979). In the first phase of the D'Agostino and Pittman (1982) study, participants' sense of prediction and control was undermined by providing noncontingent feedback as participants tried to discover a rule to describe a series of numerals. In the second phase, in an ostensibly unrelated experiment, participants per-

[7] Consistent with numerous person memory studies (e.g., Hastie & Kumar, 1979; Srull, 1981), neutral information in the present study was remembered significantly less well than either consistent or inconsistent information, making it an inappropriate baseline comparison for testing for approach versus avoidance.

formed an estimation task in which they were instructed to estimate the proportion of trials in which they had control over the presented stimuli. Participants were permitted to take as many trials as they wished before making their estimate. Presumably, being accurate on such a task should be of particular importance to control-deprived people because detecting patterns of covariation is a key component of prediction and control (Anderson, 1995; Ji, Peng, & Nisbett, 2000). Accordingly, D'Agostino and Pittman found that control-deprived participants took significantly more trials before rendering their judgment compared with participants in the control group. According to D'Agostino and Pittman's account, the experience of control deprivation led those participants to become more methodical and systematic in their information gathering as a means of restoring their sense of epistemic mastery over the environment.

In the present experiment, we attempted to undermine some subjects' subjective sense of prediction and control by presenting these subjects with theory-violating behavior and then assessing (a) their change in affective state and (b) their persistence on a subsequent control-related task. If theory violation operates similarly to noncontingent feedback, namely, by undermining one's sense of prediction and control, then when incremental theorists, for example, learn that a math geek, after taking the rigorous remedial course, still scored a 460 on the Verbal GRE (reflecting an inability to learn new skills), they should experience a rise in anxiety that will drive them to take more trials before rendering their estimate of control. Similarly, when entity theorists learn that a math geek scored poorly on a Math GRE (reflecting inconsistency in his core nature), they should experience an analogous rise in anxiety that would mediate their increased effort on the control estimation task.

Whereas several of the studies examining the effects of control deprivation have required participants to perform person–situation attribution tasks (e.g., Pittman & Pittman, 1980; Edwards & Weary, 1993), it was essential that the present task not involve these kinds of attribution processes because of previous research demonstrating clear differences between entity and incremental theorists on such tasks (e.g., Chiu et al., 1997; Molden, Plaks, & Dweck, 2004b). We did not have any a priori reason to expect entity and incremental theorists to differ on the control estimation task, because it did not concern person perception.

Method

Participants. Eighty-four University of Washington undergraduates (53 women and 31 men) provided informed consent and participated in exchange for extra course credit.

Procedure. The Implicit Person Theories Measure—Intelligence Version, used to classify participants as entity and incremental theorists of intelligence, was administered in a battery session 2–4 weeks prior to the experimental session. In the experimental session, participants first completed a questionnaire assessing their current emotional state (Higgins, 1987). The purpose of this questionnaire, which assessed participants' current affective level on two separate dimensions (anxiety–relaxation, sadness–joy), was to obtain a pretask anxiety baseline with which we could later compare participants' posttask anxiety. Of primary interest was participants' response to the question, "How anxious are you feeling right now?"

Next, participants read a brief vignette about "Brad," a college student and an "arts/humanities person." (Experiments 1 and 2 revealed that whether participants were told that Brad was a math/sciences person or an arts/humanities person did not influence their pattern of responding. Therefore, to simplify the design, this experiment included only an arts/humanities target condition.) Participants read that Brad's university, in the interest of producing well-rounded graduates, required entering students with subpar backgrounds in mathematics to undergo an intensive course in calculus and statistics to provide focused instruction to improve their weaknesses. They also read that Brad looked forward to the course and tried his best.

Next, participants read a single target behavior. This behavior was manipulated between subjects. Participants in the consistent–trait and inconsistent–trait conditions read that Brad had finished the course and decided to take a practice GRE exam. Those in the consistent–trait condition read, "On this exam, he gets a 740 on the Verbal section and a 480 on the Math section" (a consistent–trait outcome because the target appeared to remain essentially unchanged from his nature as a humanities person, showing an inability to cultivate mathematics skills). Those in the inconsistent–trait condition read, "On this exam, he gets a 480 on the Verbal section and a 740 on the Math section" (an inconsistent–trait outcome because the target appeared to change radically from his previous nature as a humanities person). Participants in the consistent–associate and inconsistent–associate conditions read, "Brad submits a poem to the University's literary magazine and it is accepted for publication" (a consistent–associate outcome because although such a behavior is highly typical of a humanities person, it was not rated as violating the basic nature of a math/sciences person). Inconsistent–associate participants read, "Brad joins the Engineering Department's team that builds and races solar-powered cars" (an inconsistent–associate outcome because although such a behavior is highly typical of a math/sciences person, it was not rated as violating the basic nature of a humanities person). After reading one of these four types of target behavior, participants completed the affect measure again, to assess posttask affect.

Participants were then seated at a computer for what was ostensibly an unrelated task by another experimenter. Instructions noted that on each trial, they would see a row of *A*s on the screen. On each trial, after a "3 . . . 2 . . . 1" countdown, participants were to press or not press the space bar and then observe what happened. Following participants' response (or nonresponse), the row of *A*s would either turn into a row of *B*s or remain a row of *A*s. Their task was to determine the percentage of trials on which they had control over whether the *A*s turned into *B*s. (Unknown to participants, the *B*s were activated by the program according to a predetermined schedule that yielded 35% control; D'Agostino & Pittman, 1982.) Participants were instructed to sample both the "press" and "not-press" responses frequently. They were further instructed that they could take as few or as many trials as they needed in order to make their judgment. Once they felt they had collected enough information, they could press the Esc key to quit the program. The principal dependent measure in this experiment was the number of trials taken by each participant before quitting. In accord with D'Agostino and Pittman (1982), we assumed that a greater number of trials reflected more methodical and systematic effort to assess one's level of control.

Following the last trial, all participants indicated their judgment of control on a scale from 0% (*no control*) to 100% (*complete control*). Because the computer was set to a known level of control (35%), we could assess the accuracy of participants' estimates. As in typical control deprivation studies (e.g., D'Agostino & Pittman, 1982; Pittman & Pittman, 1980), this was done to rule out a potential alternate explanation for observed effects: If control-deprived participants were less accurate in their assessments of control, even after extra trials, it would be possible to argue that rather than inducing differential motivation, control deprivation instead induces a cognitive deficit. If control-deprived participants' estimates were as accurate as those of non-control-deprived participants, this explanation becomes less plausible. Following the completion of the experimental procedure, participants were fully debriefed.

Results and Discussion

Implicit Person Theories Measure. Participants' responses to the Implicit Person Theories Measure items were highly reliable (Cronbach's $\alpha = .90$). As in the previous experiments, participants with a mean theory score of 3.0 or below (indicating overall agreement) were classified as entity theorists ($n = 33$), and participants with mean scores of 4.0 and above (indicating overall disagreement) were classified as incremental theorists ($n = 36$). Participants with mean theory scores that fell between 3.0 and 4.0 were unclassified ($n = 15$) and were excluded from the analyses.

Number of trials taken. The data of 4 outlying participants who either quit the task immediately or took a number of trials more than 3 standard deviations higher than the mean were excluded. To test whether theory violation would lead to intensified effort on the control estimation task, the total number of trials that each participant performed was submitted to a 2 (theory: entity vs. incremental) × 2 (consistency: consistent vs. inconsistent) × 2 (dimension: trait vs. associate) ANOVA. (Note that unlike in the previous studies, the consistency and dimension variables were between-subjects variables.) This analysis revealed a significant Consistency × Theory interaction, $F(1, 59) = 4.89$, $p < .03$, that was qualified by the predicted three-way interaction, $F(1, 59) = 4.30$, $p < .05$.

As predicted, entity and incremental theorists differed significantly in their responses following the presentation of consistent–trait versus inconsistent–trait behaviors, $F(1, 59) = 5.22$, $p < .03$. As indicated in Table 3, entity theorists took more trials after being presented with inconsistent–trait behavior ($M = 35.13$) than with consistent–trait behavior ($M = 24.62$), $F(1, 59) = 4.35$, $p < .05$, but incremental theorists showed the reverse, taking more trials after being presented with consistent–trait behavior ($M = 30.50$) than with inconsistent–trait behavior ($M = 20.38$), $F(1, 59) = 4.86$, $p < .04$. In contrast, and as predicted, entity and incremental theorists took an equivalent number of trials following the presentation of consistent–associate and inconsistent–associate behaviors, $F(1, 59) < 0.50$, *ns.* In other words, entity and incremental theorists both appeared to become more deliberate and methodical after being presented with theory-violating information than after being presented with information that either confirmed or was irrelevant to their respective theories.

Table 3
Mean Number of Trials Taken and Change in Anxiety as a Function of Type of Information and Participants' Implicit Theory, Experiment 3

	Type of information			
	Trait dimension		Associate dimension	
Theory	Consistent	Inconsistent	Consistent	Inconsistent
Entity theorists				
Trials taken				
M	24.62	35.13	28.33	24.88
SD	14.91	7.36	9.01	9.88
Anxiety change				
M	−1.00	2.13	−0.78	−0.50
SD	1.41	0.99	1.30	0.84
Incremental theorists				
Trials taken				
M	30.50	20.38	24.13	20.00
SD	8.37	7.23	11.44	9.70
Anxiety change				
M	1.92	−0.63	0.00	−0.25
SD	1.44	0.52	0.53	1.17

Note. Positive anxiety change scores indicate increased anxiety; negative scores indicate decreased anxiety.

In addition, although subjects generally tended to overestimate their degree of control ($M = 55\%$ vs. the preset level of 35%), a 2 (information consistency: consistent vs. inconsistent) × 2 (information dimension: trait vs. associate) × 2 (theory: entity vs. incremental) between-subjects ANOVA revealed no reliable differences (all Fs < 1.47, all ps $> .23$). This indicated that (a) entity and incremental theorists did not differ a priori in their estimations of control and (b) theory violation did not induce a cognitive deficit that impaired participants' ability to detect covariation.

Affect data. To test whether theory violation would lead to increased self-reported anxiety, we first calculated each participant's change-in-anxiety score (posttask minus pretask, such that larger values indicated an increase in anxiety and negative scores indicated a decrease). (Entity and incremental theorists did not differ significantly in pretask anxiety, $t < 1.3$, *ns.*) The change-in-anxiety values were submitted to a 2 (theory: entity vs. incremental) × 2 (consistency: consistent vs. inconsistent) × 2 (dimension: trait vs. associate) between-subjects ANOVA. (Two participants did not complete the affect measure.) As with the control estimation task above, the predicted Theory × Consistency × Dimension interaction was found, $F(1, 57) = 21.29$, $p < .01$.[8]

As predicted, entity and incremental theorists displayed different patterns of anxiety following exposure to trait dimension behavior, $F(1, 57) = 26.88$, $p < .01$. As depicted in Table 3, entity theorists experienced a significant increase in anxiety when faced with inconsistent–trait information ($M = 2.13$) versus consistent–trait information ($M = -1.00$), $F(1, 57) = 30.07$, $p < .01$; incremental theorists' increase in anxiety was significantly greater when faced with consistent–trait information ($M = 1.92$) than when faced with inconsistent–trait information ($M = -0.63$), $F(1, 57) = 23.87$, $p < .01$. As predicted, neither entity nor incremental theorists exhibited a significant change in anxiety following exposure to consistent–associate or inconsistent–associate behavior, $F(1, 57) < 1.0$, *ns.* In addition, neither group exhibited a significant change in sadness, depression, or any of the positive emotions on the measure (e.g., "happy," "elated") as a function of type of information (all Fs < 1.0).

Mediational analyses. Although participants' pattern of trials taken and anxiety are consistent with our hypotheses, they do not indicate whether the difference in trials taken following consistent–trait versus inconsistent–trait behaviors was mediated by change in anxiety. To test this mediational model, regression analyses were conducted following the steps outlined by Baron and Kenny (1986).

First, because consistent information on the trait dimension violates the incremental theory, and inconsistent information vio-

[8] This analysis also revealed less theoretically meaningful effects for dimension, $F(1, 59) = 12.54$, $p < .01$, and Theory × Consistency, $F(1, 59) = 30.93$.

lates the entity theory, we created a dichotomous variable, information type, that was applied to all participants, with 0 representing theory nonviolating and 1 representing theory violating. In Step 1, a regression analysis with trials as the dependent variable and information type as the independent variable revealed a significant effect ($\beta = .46$, $p < .01$).[9] In Step 2, a regression analysis with anxiety change as the dependent variable and information type as the independent variable also revealed a significant effect ($\beta = .78$, $p < .01$). In Step 3, a regression analysis with trials as the dependent variable and anxiety change and information type as the independent variables revealed a near significant relationship between anxiety change and trials ($\beta = .41$, $p = .09$), but the path between information type and trials was no longer significant ($\beta = .14$, *ns*; Sobel test: $Z = 1.70$, $p < .09$; simple bivariate relationship between anxiety charge and trials = 0.38, $p < .001$). That is, as indicated in Figure 2, for behaviors on the trait dimension, the significant relationship between information type (theory-consistent vs. theory-inconsistent) and number of trials taken was reduced when a path for change in anxiety was included in the model. Analogous analyses conducted with associate dimension behaviors revealed no significant relationships for any of the paths.

This pattern is consistent with the hypothesis that anxiety plays a mediating role in the relationship between exposure to theory-violating information and increased effort to restore the subjective sense of prediction and control. These data suggest that the experience of theory violation induces anxiety that people, if given the opportunity, strive to reduce through more methodical processing aimed at bolstering their sense of being competent detectors of covariation (D'Agostino & Pittman, 1982).

Other possible means of coping with theory violation. Previous research has indicated that if participants feel that enhanced effort will not help to improve performance, or if failure can be attributed to external factors, then control deprivation is less likely to yield intensified information gathering (D'Agostino & Pittman, 1982; Snyder, Stephan, & Rosenfield, 1976). Presumably, in the present study neither of these circumstances occurred, although these variables could be manipulated in future studies. In a related vein, intensified effort on the task is likely not the only option available to people. In recent research, it has been shown that when people had been previously given the opportunity to express or affirm a valued aspect of the self, they subsequently displayed less defensive processing of undesired (D. K. Sherman, Nelson, & Steele, 2000) and counterattitudinal (Cohen, Aronson, & Steele, 2000) information. Would the opportunity to affirm a valued aspect of the self attenuate people's apparently powerful desire to reestablish the sense of prediction mastery? It is plausible that general self-affirmation and the particular motivation to reestablish a sense of prediction mastery are in many respects intertwined and may both have contributed to participants' behavior in Experiment 3. The relative contribution of these two processes is an empirical question for future research.

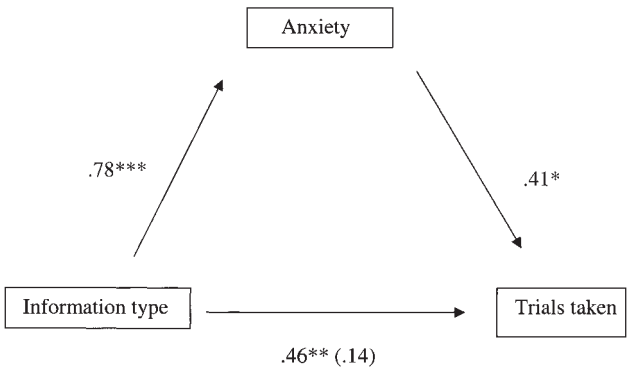

Figure 2. The relationship between information type, anxiety, and number of trials (Experiment 3). *$p < .10$. **$p < .01$. ***$p < .001$.

In sum, Experiment 3 provided evidence that unavoidable exposure to theory-violating information induced in both entity and incremental theorists an increase in anxiety and a concomitant redoubling of information-gathering efforts on a subsequent, control-relevant task (Alloy & Abramson, 1979; D'Agostino & Pittman, 1982). These data begin to illustrate how cognitive, affective, and motivational forces work together to restore the sense of prediction and control following the experience of theory violation.

General Discussion

In Experiment 1, we provided evidence consistent with our central hypothesis: For entity theorists under high cognitive load, counterstereotypic information that also violated their theory triggered more defensive processing than information that was "merely" counterstereotypic. On the other hand, for high load incremental theorists, stereotypic behavior that implied an inability to change triggered more defensive processing than behavior that was "merely" stereotypic. As described, this finding has noteworthy implications for the person memory literature because it suggests that when perceivers allocate attention to a given target behavior, they go beyond its stereotype consistency and also consider its theory consistency. In Experiment 2, by greatly varying the processing conditions, we demonstrated that people tend to follow more passive forms of motivated cognition (e.g., selective attention) when under high cognitive load but more active forms of motivated cognition (e.g., intensified scrutiny) when plentiful cognitive resources are available. In Experiment 3, we provided evidence for a mediational path between violation of one's theory, increased anxiety, and efforts to restore the perceived sense of prediction and control. To our knowledge, these experiments represent the first systematic examination of the social–cognitive mechanisms through which people protect their implicit theories of personality from invalidating information.

The Specific Affective Consequences of Theory Violation

It is of note that in Experiment 3, there was no observed relationship between attention to theory-violating information and experiencing sadness or depression. Thus, at least in the short term, the negative affective consequences of exposure to theory-violating information appear to be limited to an increase in anxiety. This is consistent with findings in the cognitive dissonance liter-

[9] Initial regression analyses that included the variables theory and Theory × Trait (consistent vs. inconsistent) revealed that these variables contributed negligibly to the overall variance, indicating that the effects were essentially the same for entity and incremental theorists. Thus, the regression analyses reported are collapsed across these variables.

ature (e.g., Elliot & Devine, 1994), in which the induction of dissonance has led to an increase in anxiety-related emotions such as discomfort and agitation but not in depression-related emotions such as sadness. On the other hand, a repeated history of exposure to theory-violating information may lead to depression in the long run. This would be consistent with prior research demonstrating the relationship between a reduced sense of prediction and control and chronic depression (e.g., Weary, Gleicher, & Marsh, 1993). It would be interesting in future research to investigate the affective and cognitive consequences of repeated theory violation over time.

The Absence of Approach Effects

In Experiment 1, when neutral (theory-irrelevant) behaviors were taken into account in the analyses, participants displayed more evidence of avoiding theory-violating information than approaching theory-confirming information. What is the reason for this asymmetry? Although people may prefer theory-confirming information to theory-violating information, if their theory is not being threatened, people may find the simple fact that their theory was confirmed comparatively uninteresting. Making reasonably accurate predictions is, after all, the theory's job. Theory-violating information, on the other hand, reflects badly on the theory's predictive accuracy and should therefore be disturbing. Consider, for example, that a person does not become elated each time he or she flips the light switch and the light comes on, but when the switch is flipped and the room stays dark, this violation of the expected order is troubling (over and above the fact that the person is standing in the dark). This may explain why in Experiment 1 (as well as in Plaks, Grant, & Dweck, 2004), participants avoided theory-violating information more assiduously than they approached theory-confirming information.

When might perceivers be more likely to approach theory-confirming information? Recall that in Experiment 3, an episode of emphatic theory violation precipitated efforts to restore control in a new domain. On the basis of this finding, one might hypothesize that following an episode of theory violation, if given the opportunity, both entity theorists and incremental theorists will be more likely to approach theory-confirming information as a means of restoring a sense of prediction and control. This is a hypothesis for future research.

Theory Violation Versus Theory Replacement

Although the present experiments have provided evidence that people are reluctant to abandon their active implicit theory, even to the point of using biased processing, other recent studies have shown that people can be readily persuaded to adopt a contrary theory. For example, experiments have shown that both the entity and incremental theories can be situationally primed (regardless of participants' chronic theory) using mock scientific articles describing evidence that human personality and behavior is fixed or malleable (Chiu et al., 1997, Experiment 5; Hong, Chiu, Dweck, Lin, & Wan, 1999, Experiment 3; Levy, Stroessner, & Dweck, 1998, Experiment 4; Plaks et al., 2001, Experiment 3). Is this a contradiction?

We believe not, because of the key distinction between theory violation and theory replacement. In the present experiments, when participants were faced with theory-violating information, no suitable replacement theory was offered. In the theory-priming studies, however, a coherent, "scientifically validated" theory was described in the article participants read. It appears, then, that people's foremost aim may be to avoid being left "theoryless." When faced with theory-violating information only—and no viable alternative theory—people engage in biased processing to protect their theory. However, when provided with a viable alternative theory, people seem more willing to accept the new alternative. In this respect, lay theorists may behave like scientific theorists. Kuhn (1962), in his analysis of the process of a scientific revolution, proposed that scientists working in a prevailing paradigm at first reject or dismiss contradictory data. Only when a coherent and intuitive alternative theory emerges—one that accounts for more of the known data—do scientists begin to shift to the new paradigm.

Implications for Stereotyping and Prejudice

The distinction between theory violation and theory replacement implies that simply exposing people to counterstereotypic information is likely to have limited effectiveness at reducing stereotyping. After all, the present experiments suggest that if the counterstereotypic information violates the perceiver's theory (as it often, though not always, does in the case of entity theorists, who typically represent about 40% of the population), motivated strategies may be used to blunt its impact. A more effective approach may be to provide an alternate mental model for understanding the group in question, or perhaps an alternate understanding of human personality and behavior in general. Accordingly, Levy and colleagues have contrasted the generally poor results yielded by simply providing counterstereotypic information with the finding that teaching the incremental theory to people reduces their propensity for stereotyping (e.g., Levy & Dweck, 1998, 1999). By changing people's meaning system so that it allows for malleability and variability of behavior, a significant reduction in stereotypic thought was produced.

Future Extensions

If, as we have argued, implicit theories truly function as frameworks through which people gauge their ability to understand the world, then motivated theory-protection effects should be evident for other beliefs, theories, and assumptions that play a pivotal role in structuring one's understanding of the social world (e.g., theories about free will vs. determinism; "just world" theories; implicit theories of relationships). That is, whenever there is a threat to a core implicit theory used for imparting meaning to a complex world, people should engage in defensive processing aimed at restoring the integrity of their lay theoretical framework.

Moreover, such effects should occur beyond the person memory context used in the present studies. For example, people may be sensitive to whether their own behavior confirms or violates their implicit theory. Adopting an implicit theories approach may provide an important extension to self-verification theory (Swann, 1990; Swann, Stein-Seroussi, & Giesler, 1992) by identifying specific self-beliefs—beyond a generalized sense of positivity or negativity—that people are highly motivated to verify. We are currently investigating these hypotheses.

Conclusion

Implicit theories generate concrete predictions about how human beings tend to act, but as we have shown in this article, these predictions are not always tested in an optimal manner. Perceivers appear to engage in motivated processing distortions that favor the preservation of their working theory. This tendency, we argue, has important implications for understanding how people mentally represent and process incoming social information and may shed light on key issues in the literatures on person memory, attribution, and motivated cognition.

References

Abramson, L. Y., Seligman, M. E., & Teasdale, J. D. (1978). Learned helplessness in humans: Critique and reformulation. *Journal of Abnormal Psychology, 87,* 49–74.

Alloy, L. B., & Abramson, L. Y. (1979). Judgment of contingency in depressed and nondepressed students: Sadder but wiser? *Journal of Experimental Psychology: General, 108,* 441–485.

Altemeyer, B. (1988). *Enemies of freedom: Understanding right-wing authoritarianism.* San Francisco: Jossey-Bass/Pfeiffer.

Anderson, C. A. (1995). Implicit personality theories and empirical data: Biased assimilation, belief perseverance and change, and covariation detection sensitivity. *Social Cognition, 13,* 25–48.

Baron, R. M., & Kenny, D. A. (1986). The moderator–mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology, 51,* 1173–1182.

Bartholow, B. D., Fabiani, M., Gratton, G., & Bettencourt, B. A. (2001). A psychophysiological examination of cognitive processing of and affective responses to social expectancy violations. *Psychological Science, 12,* 197–204.

Baumeister, R., Dale, K., & Sommer, K. L. (1998). Freudian defense mechanisms and empirical findings in modern social psychology: Reaction formation, projection, displacement, undoing, isolation, sublimation, and denial. *Journal of Personality, 66,* 1081–1124.

Bodenhausen, G. V. (1988). Stereotypic biases in social decision making and memory: Testing process models of stereotype use. *Journal of Personality and Social Psychology, 55,* 726–737.

Bonanno, G. A., Davis, P. J., Singer, J. L., & Schwartz, G. E. (1991). The repressor personality and avoidant information processing: A dichotic listening study. *Journal of Research in Personality, 25,* 386–401.

Burger, J. M., & Hemans, L. T. (1988). Desire for control and the use of attribution processes. *Journal of Personality, 56,* 531–546.

Chaiken, S., Giner-Sorolla, R., & Chen, S. (1996). Beyond accuracy: Defense and impression motives in heuristic and systematic information processing. In P. M. Gollwitzer & J. A. Bargh (Eds.), *The psychology of action: Linking cognition and motivation to behavior* (pp. 553–578). New York: Guilford Press.

Chiu, C. (1994). *Bases of categorization and person cognition.* Unpublished doctoral dissertation, Columbia University.

Chiu, C., Hong, Y., & Dweck, C. S. (1997). Lay dispositionism and implicit theories of personality. *Journal of Personality and Social Psychology, 42,* 116–131.

Cohen, G. L., Aronson, J., & Steele, C. M. (2000). When beliefs yield to evidence: Reducing biased evaluation by affirming the self. *Personality and Social Psychology Bulletin, 26,* 1151–1164.

D'Agostino, P. R., & Pittman, T. S. (1982). Effort expenditure following control deprivation. *Bulletin of the Psychonomic Society, 19,* 282–283.

Dijksterhuis, A., van Knippenberg, A., Kruglanski, A. W., & Schaper, C. (1996). Motivated social cognition: Need for closure effects on memory and judgment. *Journal of Experimental Social Psychology, 32,* 254–270.

Ditto, P. H., Scepansky, J. A., Munro, G. D., Apanovitch, A. M., & Lockhart, L. K. (1998). Motivated sensitivity to preference-inconsistent information. *Journal of Personality and Social Psychology, 75,* 53–69.

Driscoll, D. M., Hamilton, D. L., & Sorrentino, R. L. (1991). Uncertainty orientation and recall of person-descriptive information. *Personality and Social Psychology Bulletin, 17,* 494–500.

Dweck, C. S. (1999). *Self-theories: Their role in motivation, personality, and development.* Philadelphia: Psychology Press.

Dweck, C. S., Chiu, C., & Hong, Y. (1995). Implicit theories and their role in judgments and reactions: A world from two perspectives. *Psychological Inquiry, 6,* 267–285.

Dweck, C. S., & Leggett, E. L. (1988). A social–cognitive approach to motivation and personality. *Psychological Review, 25,* 109–116.

Eagly, A. H., Chen, S., Chaiken, S., & Shaw-Barnes, K. (1999). The impact of attitudes on memory: An affair to remember. *Psychological Bulletin, 125,* 65–89.

Eagly, A. H., Kulesa, P., Chen, S., & Chaiken, S. (2000). Why counterattitudinal messages are as memorable as proattitudinal messages: The importance of active defense against attack. *Personality and Social Psychology Bulletin, 26,* 1392–1408.

Edwards, J. A., & Weary, G. (1993). Depression and the impression formation continuum: Piecemeal processing despite the availability of category information. *Journal of Personality and Social Psychology, 64,* 636–645.

Elliot, A. J., & Devine, P. G. (1994). On the motivational nature of cognitive dissonance: Dissonance as psychological discomfort. *Journal of Personality and Social Psychology, 67,* 382–394.

Erdley, C. A., & Dweck, C. S. (1993). Children's implicit personality theories as predictors of their social judgments. *Child Development, 64,* 863–878.

Fletcher, G. J. O., Danilovics, P., Fernandez, G., Peterson, D., & Reeder, G. D. (1986). Attributional complexity: An individual differences measure. *Journal of Personality and Social Psychology, 51,* 875–884.

Förster, J., Higgins, E. T., & Strack, F. (2000). When stereotype disconfirmation is a personal threat: How prejudice and prevention focus moderate incongruency effects. *Social Cognition, 18,* 178–197.

Frey, D. (1986). Recent research on selective exposure to information. *Advances in Experimental Social Psychology, 19,* 41–80.

Gleicher, F., & Weary, G. (1991). The effect of depression on the quantity and quality of social inferences. *Journal of Personality and Social Psychology, 61,* 105–114.

Hashtroudi, S., Mutter, S. A., Cole, E. A., & Green, S. K. (1984). Schema-consistent and schema-inconsistent information: Processing demands. *Personality and Social Psychology Bulletin, 10,* 269–278.

Hastie, R. (1984). Causes and effects of causal attribution. *Journal of Personality and Social Psychology, 46,* 44–56.

Hastie, R., & Kumar, P. A. (1979). Person memory: Personality traits as organizing principles in memory for behaviors. *Journal of Personality and Social Psychology, 37,* 25–38.

Higgins, E. T. (1987). Self-discrepancy: A theory relating self and affect. *Psychological Review, 94,* 319–340.

Hong, Y. (1994). *Predicting trait versus process inferences: The role of implicit theories.* Unpublished doctoral dissertation, Columbia University.

Hong, Y., Chiu, C., Dweck, C. S., Lin, D. M., & Wan, W. (1999). Implicit theories, attributions, and coping. *Journal of Personality and Social Psychology, 77,* 588–599.

Hong, Y., Levy, S. R., & Chiu, C. (2001). The contribution of the lay theories approach to the study of groups. *Personality and Social Psychology Review, 5,* 98–106.

Janoff-Bulman, R. (1992). *Shattered assumptions: Towards a new psychology of trauma.* New York: Free Press.

Ji, L.-J., Peng, K., & Nisbett, R. E. (2000). Culture, control, and perception of relationships in the environment. *Journal of Personality and Social Psychology, 78,* 943–955.

Johnson, J. A., Gerner, C. K., Efran, J. S., & Overton, W. F. (1988). Personality as the basis for theoretical predilections. *Journal of Personality and Social Psychology, 55,* 824–835.

Johnston, W. A., Hawley, K. J., Plewe, S. H., Elliott, J. M. G., & DeWitt, M. J. (1990). Attention capture by novel stimuli. *Journal of Experimental Psychology: General, 119,* 397–411.

Judd, C. M., Kenny, D. A., & McClelland, G. H. (2001). Estimating and testing mediation and moderation in within-subject designs. *Psychological Methods, 6,* 115–134.

Kelley, H. H. (1967). Attribution theory in social psychology. In D. Levine (Ed.), *Nebraska Symposium on Motivation* (Vol. 15, pp. 192–241). Lincoln: University of Nebraska Press.

Kelly, G. A. (1955). *A psychology of personal constructs.* Oxford, England: Norton.

Kim, N., & Ahn, W. (2002). Clinical psychologists' theory-based mental representations of mental disorders predict their diagnostic reasoning and memory. *Journal of Experimental Psychology: General, 131,* 451–476.

Kuhn, T. S. (1962). *The structure of scientific revolutions.* Chicago: University of Chicago Press.

Levy, S. R., & Dweck, C. S. (1998). Trait- versus process-focused social judgment. *Social Cognition, 16,* 151–172.

Levy, S. R., & Dweck, C. S. (1999). The impact of children's static versus dynamic conceptions of people on stereotype formation. *Child Development, 70,* 1163–1180.

Levy, S. R., Plaks, J. E., Hong, Y., Chiu, C., & Dweck, C. S. (2001). Static versus dynamic theories and the perception of groups: Different routes to different destinations. *Personality and Social Psychology Review, 5,* 156–168.

Levy, S. R., Stroessner, S. J., & Dweck, C. S. (1998). Stereotype formation and endorsement: The role of implicit theories. *Journal of Personality and Social Psychology, 74,* 1421–1436.

MacLeod, C., Matthews, A., & Tata, P. (1986). Attentional bias in emotional disorders. *Journal of Abnormal Psychology, 95,* 15–20.

Macmillan, N. A., & Creelman, C. D. (1991). *Detection theory: A user's guide.* New York: Cambridge University Press.

Macrae, C. N., Bodenhausen, G. V., Schloerscheidt, A. M., & Milne, A. B. (1999). Tales of the unexpected: Executive function and person perception. *Journal of Personality and Social Psychology, 76,* 200–213.

McConnell, A. R. (2001). Implicit theories: Consequences for social judgment of individuals. *Journal of Experimental Social Psychology, 37,* 215–227.

McGregor, I., & Marigold, D. C. (2003). Defensive zeal and the uncertain self: What makes you so sure? *Journal of Personality and Social Psychology, 85,* 838–852.

Molden, D. C., Plaks, J. E., & Dweck, C. S. (2004a). *A meaning systems approach to predictions of P × S consistency*. Manuscript in preparation.

Molden, D. C., Plaks, J. E., & Dweck, C. S. (2004b). *A meaning systems approach to social inference: Lay theories of personality and inferential processes.* Manuscript under review.

Morris, M. W., Ames, D. R., & Knowles, E. D. (2001). What we theorize when we theorize that we theorize: Examining the "implicit theory" construct from a cross-disciplinary perspective. In G. Moskowitz (Ed.), *Cognitive social psychology: The Princeton Symposium on the Legacy and Future of Social Cognition* (pp. 143–161). Mahwah, NJ: Erlbaum.

Murphy, G. L., & Medin, D. L. (1985). The role of theories in conceptual coherence. *Psychological Review, 92,* 289–316.

Newman, L. S., & McKinney, L. C. (2002). Repressive coping and threat avoidance: An idiographic Stroop study. *Personality and Social Psychology Bulletin, 28,* 409–422.

Olson, J. M., Roese, N. J., & Zanna, M. P. (1996). Expectancies. In E. T. Higgins & A. W. Kruglanski (Eds.), *Social psychology: Handbook of basic principles* (pp. 211–238). New York: Guilford Press.

Pepper, S. (1942). *World hypotheses.* Berkeley: University of California Press.

Pittman, T. S., & D'Agostino, P. R. (1989). Motivation and cognition: Control deprivation and the nature of subsequent information processing. *Journal of Experimental Social Psychology, 25,* 465–480.

Pittman, T. S., & Pittman, N. L. (1980). Deprivation of control and the attribution process. *Journal of Personality and Social Psychology, 39,* 377–389.

Plaks, J. E. (2001). *Implicit theories and attention allocation: Preferences for stereotypic vs. counterstereotypic information.* Unpublished doctoral dissertation, Columbia University.

Plaks, J. E., Grant, H., & Dweck, C. S. (2004). [Implicit theories and selective attention to trait versus associate stereotype-relevant information.] Unpublished raw data, University of Washington.

Plaks, J. E., Levy, S. R., Dweck, C. S., & Stroessner, S. J. (2004). In the eye of the beholder: Lay theories and the perception of group entitativity, variability, and essence. In V. Yzerbyt, C. M. Judd, & O. Corneille (Eds.), *The psychology of group perception: Perceived variability, entitativity, and essentialism* (pp. 127–146). New York: Psychology Press.

Plaks, J. E., Stroessner, S. J., Dweck, C. S., & Sherman, J. W. (2001). Person theories and attention allocation: Preferences for stereotypic versus counterstereotypic information. *Journal of Personality and Social Psychology, 80,* 876–893.

Sherman, D. K., Nelson, L. D., & Steele, C. M. (2000). Do messages about health risks threaten the self? Increasing the acceptance of threatening health messages via self-affirmation. *Personality and Social Psychology Bulletin, 26,* 1046–1058.

Sherman, J. W., Conrey, F. R., & Groom, C. J. (2004). Encoding flexibility revisited: Evidence for enhanced encoding of stereotype-inconsistent information under cognitive load. *Social Cognition, 22,* 214–232.

Sherman, J. W., & Frost, L. A. (2000). On the encoding of stereotype-relevant information under cognitive load. *Personality and Social Psychology Bulletin, 34,* 26–34.

Sherman, J. W., Lee, A. Y., Bessenoff, G. R., & Frost, L. A. (1998). Stereotype efficiency reconsidered: Encoding flexibility under cognitive load. *Journal of Personality and Social Psychology, 75,* 589–606.

Skowronski, J. J. (2002). Honesty and intelligence judgments of individuals and groups: The effects of entity-related behavior diagnosticity and implicit theories. *Social Cognition, 20,* 136–169.

Skowronski, J. J., & Carlston, D. E. (1987). Social judgment and social memory: The role of cue diagnosticity in negativity, positivity, and extremity biases. *Journal of Personality and Social Psychology, 52,* 689–699.

Snyder, M. L., Stephan, W. G., & Rosenfield, D. (1976). Egotism and attribution. *Journal of Personality and Social Psychology, 33,* 435–441.

Spears, R., & Haslam, S. A. (1997). Stereotyping and the burden of cognitive load. In R. Spears, P. J. Oakes, N. Ellemers, & S. A. Haslam (Eds.), *The social psychology of stereotyping and group life* (pp. 171–207). Cambridge, MA: Blackwell.

Srull, T. K. (1981). Person memory: Some tests of associative storage and retrieval models. *Journal of Experimental Psychology: Human Learning and Memory, 7,* 440–463.

Stangor, C., & Duan, C. (1991). Effects of multiple task demands upon memory for information about social groups. *Journal of Experimental Social Psychology, 27,* 357–378.

Stangor, C., & McMillan, D. (1992). Memory for expectancy-congruent and expectancy-incongruent information: A review of the social and social developmental literatures. *Psychological Bulletin, 111,* 42–61.

Swann, W. B. (1990). To be adored or to be known: The interplay of self-enhancement and self-verification. In E. T. Higgins & R. Sorrentino (Eds.), *Handbook of motivation and cognition: Vol. 2. Foundations of social behavior* (pp. 408–448). New York: Guilford Press.

Swann, W. B., Stein-Seroussi, A., & Giesler, R. B. (1992). Why people self-verify. *Journal of Personality and Social Psychology, 62,* 392–401.

Walker, A. M., & Sorrentino, R. M. (2000). Control motivation and uncertainty: Information processing or avoidance in moderate depressives and nondepressives, *Personality and Social Psychology Bulletin, 26,* 436–451.

Weary, G., Gleicher, F., & Marsh, K. (1993). *Control motivation and social cognition.* New York: Springer-Verlag.

Weary, G., Jacobson, J. A., Edwards, J. A., & Tobin, S. J. (2001). Chronic and temporarily activated causal uncertainty beliefs and stereotype usage. *Journal of Personality and Social Psychology, 81,* 206–219.

Whitehead, A. N. (1929). *Process and reality.* New York: Free Press.

Wittenbrink, B., Hilton, J. L., & Gist, P. L. (1998). In search of similarity: Stereotypes as naïve theories in social categorization. *Social Cognition, 16,* 31–55.

Zillman, D., & Zillman, M. (1996). Psychoneuroendocrinology of social behavior. In E. T. Higgins & A. W. Kruglanski (Eds.), *Social psychology: Handbook of basic principles* (pp. 39–71). New York: Guilford Press.

Received February 13, 2003
Revision received August 4, 2004
Accepted August 13, 2004 ■