Instructor: Wilson, Timothy
Course: Psyc 967
Semester_Requested: Fall '06
Email: tdw@virginia.edu
Phone: 924-0674
Author: 
Article-Chapter_Title: 
Journal-Book_Title: 
Call_Number: 
Volume: 
Date: 
Pages: 
Date_Needed_by: 07/15/06
Retrieval: The Library should retrieve this material for Toolkit
Copyright: request complies with copyright guidelines
Submit: Submit the Request

Printed for: interlibrary retrieval section on Thu, 08 Jun 2006 11:43:04 -0400
INTROSPECTION, ATTITUDE CHANGE, AND ATTITUDE-BEHAVIOR CONSISTENCY: THE DISRUPTIVE EFFECTS OF EXPLAINING WHY WE FEEL THE WAY WE DO

Timothy D. Wilson
DEPARTMENT OF PSYCHOLOGY
UNIVERSITY OF VIRGINIA
CHARLOTTESVILLE, VIRGINIA 22903

Dana S. Dunn
DEPARTMENT OF PSYCHOLOGY
MORAVIAN COLLEGE
BETHLEHEM, PENNSYLVANIA 18018

Dolores Kraft
Douglas J. Lisle
DEPARTMENT OF PSYCHOLOGY
UNIVERSITY OF VIRGINIA
CHARLOTTESVILLE, VIRGINIA 22903

1. Introduction

The ability to introspect is uniquely human. Other species may have thoughts, feelings, and motives, but as far as we know, they do not occupy their time by reflecting on the nature of these states. Because introspection is a mark of what it is to be human, and one of the highest achievements of evolution, it is tempting to view it as a generally accurate and error-free enterprise. The first experimental psychologists made this assumption by using introspective reports to study thought and perception (e.g., Titchener, 1912). Only the introspections of respondents who had undergone a good deal of training were trusted, but reports from these self-observers were the sole source of data for theories of the mind.

The use of introspective reports by the structuralists failed, due largely to the unreliability of the reports. This failure was widely criticized, and was partly responsible for the rise of behaviorism (e.g., Watson, 1913). With the waning of behaviorism and the rise of cognitive science, the use of verbal reports on attitudes, thoughts, and feelings have again become commonplace, as has a renewed debate over the validity of these reports (Ericsson & Simon, 1980, 1984; Lieberman, 1979; Nisbett & Ross, 1980; Nisbett & Wilson, 1977b; Smith & Miller, 1978; Wilson, 1985b; Wilson & Stone, 1985).
As well as being viewed as error-free, introspection has traditionally been considered to be a beneficial activity. Self-reflection is an integral part of most forms of psychotherapy, and some decision theorists have advocated increased reflection as an aid to better decision making (e.g., Janis & Mann, 1977). Many people deliberately become more reflective when faced with important decisions, and they may even make long lists outlining their thoughts and feelings about each alternative. Indeed, in this hurried, hectic world of ours, it seems that we could prevent much mindless behavior by taking the time to be more contemplative about our actions (Langer, 1978).

Recent research in social psychology has corroborated the usefulness of certain kinds of self-reflection, particularly focusing on one's attitudes and beliefs. Snyder and his colleagues, for example, have found that instructing people to think about their attitudes increased the consistency between attitudes and behavior, presumably by increasing the salience of the attitude and its implications for behavior (Snyder & Kedzierski, 1982; Snyder & Swann, 1976; see also Fazio, Chen, McDonel, & Sherman, 1982). Research generated by self-awareness theory (Duval & Wicklund, 1972; Wicklund, 1975) has also demonstrated that self-focused attention increases attitude–behavior correlations (e.g., Pryor, Gibbons, Wicklund, Fazio, & Hood, 1977; Scheier, Buss, & Buss, 1978). Finally, Tesser (1978) has shown that thinking about one's attitudes increases their strength.

Just as assumptions about the accuracy of introspection have been challenged, however, so have claims about its beneficial effects. Theodore Roethke observed that "Self-contemplation is a curse/That makes an old confusion worse" (1975, p. 249). Mario Vargas Llosa, a judge at the Berlin Film Festival, found it less than useful to introspect about why he liked or disliked films in the competition:

I went to every screening with a fresh pack of notecards that I would dutifully cover with my impressions of each and every film. The result, of course, was that the movies ceased to be fun and turned into problems, a struggle against time, darkness and my own esthetic emotions, which these autopsies confused. I was so worried about evaluating every aspect of every film that my entire system of values went into shock, and I quickly realized that I could no longer easily tell what I liked or didn't or why. (Vargas Llosa, 1986, p. 23)

In this article, we present evidence consistent with the observations of Roethke and Vargas Llosa that introspection can be disruptive. Our focus is on one type of introspection—thinking about the reasons for one's feelings. We attempt to demonstrate that this type of thought can cause people to change their minds about how they feel and lead to a disconnection between their attitudes and their behavior.

When people think about why they feel the way they do, we suggest that they feel compelled to give a "good story" to explain their feelings. The reasons they come up with emphasize cognitions about the attitude object, and the reasons are often incomplete or incorrect. Further, these cognitions often consist of a biased sample of reasons that imply a different attitude than the one subjects previously held. Unaware that their reasons are a biased sample, subjects change their attitude in the direction implied by their reasons. Because the reasons consist

of beliefs about the attributes of the attitude object, this new attitude has a heavy cognitive flavor. Behavior often remains affectionately driven, however, resulting in reduced attitude–behavior consistency.

We have chosen with care the word "disruptive" to describe the effects of thinking about reasons because this term has fewer pejorative connotations than adjectives like "harmful" or "detrimental." We present evidence suggesting that under some circumstances, people are better off not thinking about why they feel the way they do. Our argument should not be viewed, however, as a general condemnation of introspection. We do not recommend that people studiously avoid introspecting about the reasons behind their feelings. Indeed, a strength of our model is that it specifies when it is best to introspect in this manner and when it is not. Thus, we question the universal applicability of Socrates's oft-quoted statement that "the unexamined life is not worth living" (Loomis, 1942, p. 56) but we stop far short of suggesting that the examined attitude is not worth having.

II. Thinking about Reasons Reduces Attitude–Behavior Consistency

Asking people to think about reasons was initially used as a means of changing inferences about people's internal states in self-attribution experiments (Wilson, Hall, & Johnson, 1981; Wilson & Linville, 1982). Thinking about reasons was found to change people's reported attitudes, traits, and moods. In Wilson et al.'s (1981) Study 2, for example, some subjects were rewarded for playing an interesting puzzle, while others were not. Differences in reported liking for the puzzle were found only among subjects who were asked to think about why they were motivated to play with the puzzle. Among subjects who did think about reasons, an overjustification effect (Lepper & Greene, 1978) occurred. Those who were rewarded reported less interest in the puzzles than did those who were not rewarded. Among subjects who did not think about reasons, no differences in reported liking were found in the reward versus no-reward conditions.

In these early studies, thinking about reasons typically had no effect on behavioral measures of people's attitudes, such as the amount of time Wilson et al.'s (1981) subjects played with the puzzles in a free-time period. This led us to believe that thinking about reasons might adversely affect attitude–behavior consistency (the consistency question was equivocal in the early studies because the n's were too small to allow meaningful tests of differences in attitude–behavior correlations). Several subsequent studies were thus performed to test the hypothesis that thinking about reasons reduces attitude–behavior consistency.

In each experiment, subjects either had or were given direct experience with an attitude object, then their verbally reported attitudes and behaviors toward the attitude object were assessed. The attitude reports were always private and anonymous, whereas the behaviors, as far as subjects knew, were not monitored.
TABLE I (continued)

<table>
<thead>
<tr>
<th>Behavioral measure</th>
<th>Control condition</th>
<th>Reasons condition</th>
<th>z</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Playtime with puzzles</td>
<td>.54</td>
<td>.17</td>
<td>2.10</td>
<td>.04</td>
</tr>
<tr>
<td>Facial expressions</td>
<td>.57</td>
<td>-.05</td>
<td>1.77</td>
<td>.08</td>
</tr>
<tr>
<td>Facial expressions</td>
<td>.53</td>
<td>.03</td>
<td>2.06</td>
<td>.04</td>
</tr>
<tr>
<td>Persistence of relationship</td>
<td>.62</td>
<td>.10</td>
<td>2.32</td>
<td>.02</td>
</tr>
<tr>
<td>Facial expressions</td>
<td>.22</td>
<td>-.09</td>
<td>1.56</td>
<td>.12</td>
</tr>
<tr>
<td>Beverages purchased</td>
<td>.59</td>
<td>.41</td>
<td>1.61</td>
<td>.11</td>
</tr>
<tr>
<td>Playtime with puzzles</td>
<td>.53</td>
<td>.25</td>
<td>2.48</td>
<td>.01</td>
</tr>
<tr>
<td>No. of fliers taken</td>
<td>.46</td>
<td>-.43</td>
<td>2.52</td>
<td>.01</td>
</tr>
<tr>
<td>Am. of beverage consumed</td>
<td>.76</td>
<td>.22</td>
<td>2.86</td>
<td>.004</td>
</tr>
<tr>
<td>Playtime with puzzles</td>
<td>.43</td>
<td>-.02</td>
<td>2.54</td>
<td>.01</td>
</tr>
</tbody>
</table>

Wilson et al.'s (1984) Studies 2 and 3 established the generalizability of the disruptive effects of thinking about reasons. In Study 2 (actually two studies—one plus a replication), subjects viewed a series of color slides of vacation scenes. Subjects in the reasons condition were instructed to think about why they liked or disliked the pictures, while subjects in the control condition were simply told to watch the slides as if they were seeing a movie in a theater. The latter subjects exhibited a fair degree of consistency between their attitudes (reports of how much they enjoyed the slide show) and behavior (ratings made by hidden observers of the pleasantness of subjects' facial expressions during the slide show), \( r = .55 \). As predicted, subjects who thought about reasons showed no consistency, \( r = -.01 \) (these correlations are averaged across the two slide show studies).

Wilson et al.’s (1984) third study demonstrated that the effect of thinking about reasons is not limited to mundane attitudes such as those about puzzles or vacation pictures, nor is it limited to behaviors assessed in the laboratory. Students involved in steady dating relationships were either asked to list all the reasons they could think of as to why they felt the way they did about their dating partners, or were not given any instructions to introspect. Once again, thinking about reasons significantly undermined attitude–behavior consistency; in this case, the correlation was between how well-adjusted subjects said their relationship was and whether they were still going out with their dating partner several months later. This correlation was .62 in the control condition, but only .10 among those who thought about reasons.

The effects of thinking about reasons on attitude–behavior consistency has proven to be a very robust phenomenon. Table I lists 10 studies that have found...
a reduction in consistency due to a reasons analysis. Consistency has been undermined with highly self-relevant attitudes (feelings about one's dating partner) and more mundane attitudes (interest in puzzles), with a variety of behavioral measures, and in both laboratory and field situations. In addition, the effect has been replicated by other researchers (Millar & Tesser, 1986).

We should acknowledge that the list of experiments in Table 1 is selective. It does not include our initial studies that were not designed to examine differences in correlations (two studies by Wilson et al., 1981, and one by Wilson & Linville, 1982), nor does it include our failures—a few studies that found no differences between reasons and control conditions. The purpose of Table 1 is to illustrate the range of settings, attitude objects, and measures with which the phenomenon has been demonstrated. To illustrate better the robustness of the effect of thinking about reasons, we averaged across all the known studies in this area, including our initial and all subsequent studies, and our successes as well as our failures. Using the method of adding z's (Mosteller & Bush, 1954; Rosenthal, 1978), it is apparent that thinking about reasons does reduce attitude–behavior consistency, $z = 4.45, p < .00002$.

III. Why Does Thinking about Reasons Reduce Attitude–Behavior Consistency?

A. THINKING ABOUT REASONS IS DISRUPTIVE; FOCUSING ON FEELINGS IS NOT

In attempting to explain why thinking about reasons reduces attitude–behavior consistency, our first step was to compare the effects of thinking about reasons with another type of introspection, to see if the disruptive effects we found were specific to thinking about reasons. As previously mentioned, there is some evidence that it is, that several researchers have found that another form of introspection—inducing people to focus on their feelings—increases attitude–behavior consistency (Carver & Scheier, 1981; Fazio et al., 1982; Snyder, 1982; Wiekulnd, 1982).

There are several differences between our research and self-focus research, making it unclear whether the different effects that were found were due to the different forms of introspection used or to some other factor. For example, the different types of studies began with different levels of attitude–behavior consistency in the control conditions, in which people did not introspect about their beliefs and feelings. We began with attitudes about which there was high attitude–behavior consistency, and we found that thinking about reasons reduced it. The self-focus studies typically began with low consistency, and they found that focusing on one's feelings increased it. One possibility, then, is that if an attitude is weak or nonsalient (and thus does not predict behavior), then any form of introspection—be it focusing or thinking about reasons—might increase its accessibility, and improve attitude–behavior consistency. If the attitude is already accessible and predictive of behavior, then either type of introspection might call to mind competing thoughts and feelings, reducing the accessibility of the initial attitude and lowering attitude–behavior consistency.

Alternatively, the initial accessibility of the attitude may be less important than the type of introspection that is performed. Though there are surface similarities between focusing on an attitude and thinking about reasons, there are important differences between these types of introspection. The former involves a retrieval of affect, while the latter involves a retrieval not only of how one feels but also of thoughts and beliefs about those attributes of the attitude object that might explain the affective reaction. The question of why these different types of introspection can have different effects is discussed shortly. It is first necessary to establish that the discrepancy between the findings of our studies and self-focus studies is due to differences in the types of introspection people perform and not due to other differences between the studies, such as the initial accessibility of attitudes.

Wilson and Dunn (1986a) performed two studies in which focusing instructions were pitted against instructions to think about reasons, holding constant the initial accessibility of people's attitudes. There were three conditions in each study: One in which people were asked to think about why they felt the way they did about the attitude object, one in which they were asked to focus on how they felt about the attitude object (using instructions very similar to those used by Snyder & Swann, 1976), and a control condition in which subjects did not receive any introspection instructions. Study 1 was a field experiment, in which the subjects were students waiting in line for dinner at a college cafeteria, and the attitude objects were beverages served at the cafeteria. Study 2 was a replication of Wilson et al.'s (1984) puzzle experiment.

The results of the two studies are summarized in Table 1. In both studies, thinking about reasons reduced attitude–behavior correlations relative to control conditions.

<table>
<thead>
<tr>
<th>TABLE II</th>
</tr>
</thead>
<tbody>
<tr>
<td>EFFECTS OF ANALYZING REASONS Versus FOCUSING ON FEELINGS ON ATTITUDE–BEHAVIOR CONSISTENCY*</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Study</th>
<th>Focus</th>
<th>Control</th>
<th>Reasons</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wilson &amp; Dunn (1986a), Study 1</td>
<td>.63</td>
<td>.59</td>
<td>.41</td>
</tr>
<tr>
<td>Wilson &amp; Dunn (1986a), Study 2</td>
<td>.54</td>
<td>.53</td>
<td>.25</td>
</tr>
</tbody>
</table>

*The correlations were computed on a within-subject basis and then averaged after converting them to $z$-scores. The entries in the table have been converted back to correlation coefficients.
subjects, while focusing led to slight increases in correlations. Averaging across studies, the correlations in the reasons condition were significantly lower than the correlations in both the control and the focus conditions. The correlations in the focusing conditions were not significantly higher than the control correlations, possibly due to the fact that our control subjects had direct experience with the attitude objects (unlike most prior studies that used focusing manipulations), which might have caused people to focus naturally on their attitudes (Fazio & Zanna, 1981).

The important finding for our purposes was that thinking about reasons had a significantly different effect than focusing on feelings, at least when attitude–behavior correlations began at a high level. (It is possible that both thinking about reasons and focusing on feelings would increase correlations that began at a low level.) This finding has also been obtained in an experiment by Millar and Tesser (1986), which is discussed in detail later. Because the initial salience of people’s attitude was held constant in the Wilson and Dunn (1986a) and Millar and Tesser (1986) studies, it cannot be argued that either type of introspection (e.g., focusing or analyzing) reduces the accessibility of salient attitudes.

**B. THINKING ABOUT REASONS CAN CAUSE ATTITUDE CHANGE**

Thus, our first clue is that there is something about thinking about reasons—as opposed to focusing on feelings—that reduces attitude–behavior consistency. Our second clue is that thinking about reasons causes a change in people’s reported attitudes.

The lowering of a correlation between attitudes and behavior can occur for at least three reasons: (1) People could change their attitudes but not their behavior, (2) people could change their behavior but not their attitudes, or (3) people could change both their attitudes and their behavior, but in different directions. The available evidence suggests that thinking about reasons reduces attitude–behavior consistency for the first of these reasons: People change their minds about how they feel, but do not change their behavior. The effects of thinking about reasons on attitudes and behavior were examined in 20 studies that have used reasons analysis manipulations: the 10 summarized in Table I; 2 by Wilson et al., 1981; 1 by Wilson and Linville, 1982; 1 plus a replication by Tesser, Leone, and Clary, 1978; and 5 unpublished studies. Analyzing reasons changed self-reported attitudes in 11 of the 20 studies, whereas behavior changed in only 3 of the 20. (By “change,” we mean that the average reported attitude in the reasons analysis condition was significantly different than the average reported attitude in the control condition.) This summary is somewhat misleading, because it includes three studies in which the reasons-analysis manipulation failed to affect any dependent variable, including attitude–behavior consistency. When these studies are eliminated from the summary, thinking about reasons changed reported attitudes in 11 of 17, whereas behavior changed in only 3 of 17.

In addition, in four of the studies (Wilson et al.’s [1984] Studies 1 and 3, Wilson & Dunn’s [1986a] Studies 1 and 2), we coded the reasons people gave in the reasons conditions according to how positive an attitude toward the stimulus they conveyed. The positivity of subjects’ reasons correlated significantly more with their subsequent attitudes than with their behavior. That is, people’s reasons were more apt to predict the attitudes they reported than how they behaved.

Oddly, however, in some studies thinking about reasons caused a reduction in attitude–behavior correlation, in the absence of significant differences between reported attitudes or behavior in the reasons versus the control conditions. In the Wilson et al. (1984) couples study, for example, the predicted drop in correlation was found, but there were no differences in the mean reported level of adjustment of the relationship between the reasons analysis and control subjects; nor was there a difference (thankfully!) in the percentage of couples who had broken up several months later. One explanation for this pattern of results is that thinking about reasons caused people to change their minds about how they felt, but that the direction of this change differed from subject to subject. In the couples study, for example, thinking about reasons might have caused some subjects to report a more positive attitude toward their relationship and others to report a more negative attitude. Assuming that their behavior was unaffected by the reasons manipulation, this change in attitude would explain the drop in attitude–behavior consistency. The mean level of reported adjustment, however, averaged across subjects, would be indistinguishable from the mean level of reported adjustment in the control condition.

The only way to detect such a pattern of change would be to measure attitudes twice; once before a reasons analysis manipulation and once after. We did this in two recent studies, to test the possibility that thinking about reasons can cause change in a positive direction for some subjects but a negative direction for others. In the first study (Wilson, Kraft, & Dunn, 1988a), during a mass testing session, subjects rated their attitudes toward several possible candidates for President in 1988. Several weeks later, they were seen individually in the laboratory, where they rated their attitudes again toward six of these candidates. Before rating their attitudes in the lab, half of the subjects wrote down their reasons for preferring or not preferring each candidate, while the other half completed a filler questionnaire.

As in Wilson et al.’s (1984) couples study, people did not become uniformly more positive or negative toward the attitude object as a result of thinking about reasons. The average amount of change in a positive or negative direction was minimal, and it was no greater in the reasons than in the control condition. To test the possibility that the reasons manipulation caused change in a positive direction for some subjects but a negative direction for others, we computed the absolute value of the difference between each subject’s attitude at Time 1 and
Time 2. As predicted, subjects in the reasons condition showed a greater amount of absolute change than did control subjects ($M'$s = .76 and .56, respectively, $p < .05$). 1

This finding was replicated in a second study (Wilson & Kraft, 1988) in the context of a phone survey. Subjects were telephoned and asked to give their attitudes toward some social issues. Two sets of issues were used. Some subjects were asked for their attitudes toward the death penalty, a national health insurance, and busing to achieve school integration, while others were asked for their attitudes toward abortion and President Reagan. 2 Half of the subjects were first asked to give reasons for why they felt the way they did about each issue; the other half were not. All subjects had rated their attitudes toward the same six issues several weeks earlier during a mass testing session, allowing us to test the amount of attitude change that occurred in each condition. Once again, people in the reasons-analysis condition did not become uniformly more positive or negative on any of the issues, but they did exhibit significantly more change when the direction of this change is ignored.

IV. Self-Persuasion via Self-Reflection

Thinking about reasons thus appears to reduce attitude–behavior consistency by changing attitudes, while behavior remains unchanged. What remains to be explained is why thinking about reasons changes attitudes, particularly why it can cause change in a positive direction for some people and a negative direction for others. In this section, we discuss in detail the mediating factors that we propose as responsible for the disruptive effects of thinking about reasons. Though empirical support has been found for several of our proposed mechanisms, some are speculative. The processes we propose do implicate some important boundary conditions on the effects of thinking about reasons, however, and we have accumulated a good deal of evidence for one of these boundary conditions.

Our argument is as follows: First, when asked to explain their feelings, people feel compelled to come up with a reasonable-sounding answer (reasonable to themselves and to others). Second, when people attempt to explain an attitude, they do not always know exactly why they feel the way they do. Therefore, the

1These results are only for those subjects who were relatively unknowledgable about the candidates. As is discussed later, no differences were predicted for subjects who were knowledgeable, and none were found.

2These subjects were also asked their attitudes toward U.S. support for the Contras in Nicaragua. As it happened, we did our survey during the Spring of 1987, when there was an increasing amount of media attention to the diversion of funds to the Contras from arms sales to the Iranians. As a result, this was the one issue of the six we used in which control subjects showed a change in attitude between Time 1 and 2. In order to test our hypothesis that thinking about reasons causes attitude change, it is necessary that control subjects not change their attitudes. Therefore, this issue was eliminated from the final analyses.

---

THE DISRUPTIVE EFFECTS OF EXPLAINING WHY

reasons they come up with sound plausible, but they might not correctly explain their feelings, or they might only be a subset of the actual reasons underlying their attitude. If the reasons people generate are only a subsample of the correct reasons, they may be a biased sample, because what is easiest to verbalize or what is available in memory at any given time may well not be representative of the entire set of reasons. Unaware that their reasons are incomplete or incorrect, people view them as representative of their feelings and adopt the attitude they imply. In the following section, we explain this set of hypotheses in greater detail, and we suggest the way in which it accounts for attitude change in both positive and negative directions and for a reduction in attitude–behavior consistency.

A. PEOPLE ARE RARELY AT A LOSS FOR REASONS

When we ask people in our studies to explain the reasons for their attitudes, almost never do people respond by saying, "I don't know." This may, of course, be a function of the experimental context in which subjects find themselves. Subjects may not know why they feel the way they do, but they construct an explanation to avoid looking foolish or inarticulate. It is difficult to assess the extent to which people's readiness to give reasons is driven by social desirability concerns, because the best way to know if they have reasons is to ask them. We believe that there is more to this readiness than social desirability, however.

First, to minimize social desirability concerns, we tell subjects in our studies that we want them to think about their reasons in order to organize their thoughts, and we explain that no one will ever read what they write. Second, as is discussed later, we have replicated the effects of thinking about reasons on attitude–behavior consistency in studies in which subjects are never directly asked to introspect, but in which they are placed in conditions that should trigger a reasons analysis.

Even if people's readiness to give reasons in our studies is due to a desire to avoid looking foolish, it is important to note that this desire should be present in many situations other than psychology experiments. It is common in everyday life to be asked to explain one's feelings ("Why did you like the movie?"). Such questions are as likely to raise social desirability concerns in these contexts as they are in psychology experiments.

B. THE REASONS PEOPLE GIVE ARE OFTEN INCOMPLETE OR INCORRECT

People's explanations of their own feelings, judgments, and behaviors are often incorrect (Nisbett & Wilson, 1977b; Wilson & Stone, 1985). While there is some controversy over the extent of this inaccuracy (e.g., Ericsson & Simon,
were to give the correct causal account, which was that they thought he was physically attractive in part because they liked him. There appears to be a strong cultural norm in our society to explain attitudes in terms of cognitions about the attributes of the attitude object. It is much more common to hear people say, “I like the candidate because of his liberal views” than, “I like the candidate because my parents were Democrats.”

Other factors may be overrepresented in people’s accounts of their reasons because they are easy to verbalize, more available in memory, or flattering. Some factors are more difficult to verbalize or even recognize than others. A subject in Wilson et al.’s (1984) couples study, for example, might find the fact that her partner is physically attractive easier to verbalize than the fact that when she is depressed, he fails to be supportive. In addition, when asked to think of reasons, people may focus too much on stimulus characteristics that are available in memory. For example, a couple who had just had a fight might focus on the negative aspects of their relationship when asked to analyze it, when as a whole their relationship is going well.

D. PEOPLE ADOPT THE ATTITUDE IMPLIED BY THEIR BIASED SET OF REASONS

Attitude change appears to occur through a process of self-persuasion, in which people convince themselves that the biased sample of reasons they just generated is representative of their attitude, possibly due to a general insensitivity to sample bias (Hamill, Wilson, & Nisbett, 1980). Thus, the woman who finds it easier to verbalize her dating partner’s attractiveness than his lack of responsiveness may convince herself that she likes him more than she thought she did, because all of the reasons that she thinks of are positive. Similarly, the couple who just had a fight might convince themselves that their relationship is on the rocks, since all the attributes that come to their minds are negative.

Perhaps the most controversial part of our self-persuasion hypothesis is the idea that people come up with reasons that do not match their initial affect toward the attitude object. Consistency theories, as well as common sense, suggest that when asked to explain their feelings, people will search for reasons that are consistent with their affect, not inconsistent. We agree that such consistency drives can be operative; it is unlikely that a person who has extremely positive feelings toward his dating partner will come up with all negative reasons when asked to explain his feelings. Indeed, we make no claims that asking people to think about reasons will change love to hate or Democrats to Republicans. We do argue that people often equivocate about their attitudes, and that on many issues, they have a range of positions that are acceptable (C. Sherif, M. Sherif, & Nebergall, 1965; M. Sherif & Hovland, 1961). Our argument is that the attitude people have at any given moment is influenced by the reasons they bring to
mind when trying to explain it. Cognitions that were not previously a central part of the attitude may be overemphasized, leading to some attitude change. Thus, while we do not claim to be able to turn love into hate, we do claim that people’s attitudes can become somewhat more positive or negative as a result of having to explain their feelings.

We will return to this point when we discuss the boundary conditions on the effects of thinking about reasons, suggesting that people are most apt to bring to mind cognitions that conflict with their initial affect for certain kinds of attitudes. At present, it is important to note that there is evidence for the fact that attitude change can result from having a biased sample of reasons in memory (Saličnik, 1974; Seligman, Farbo, & Zanna, 1980). Seligman et al., for example, asked subjects to think of reasons why they liked their dating partner. Unlike in our studies, these researchers attempted to bias the types of reasons subjects considered by the way the question was asked: Subjects answered either the question, “I go out with this person because 1 . . . .” or, “I go out with this person in order to . . . .” The former question elicited reasons that were primarily internal (i.e., having to do with one’s own feelings and commitment), while the latter question elicited reasons that were primarily external (i.e., having to do with factors other than love and commitment, such as the desire to impress one’s friends).

As predicted, subjects seem not to have realized that they had generated a biased set of reasons, and as a result, they adopted the attitude these reasons implied. That is, subjects who answered the “because I” question reported significantly more love and expressed more of an intention to marry their partners than did subjects who answered the “in order to” question. Saličnik (1974) found similar results when the attitude object was a college course that subjects were taking.

Thus, our reasons analysis studies can be considered replications of the Saličnik (1974) and Seligman et al. (1980) studies, with two major differences. First, we did not direct subjects’ reasons in any particular direction. Interestingly, subjects seem to come up with biased samples on their own, resulting in attitude change. Second, we included behavioral measures in our studies, and we found that the new attitudes subjects reported typically did not predict their behavior.

Finally, our argument about self-persuasion via self-reflection may help explain why attitude change is not in a common direction in some of our studies. What is easiest to verbalize or most available in memory might be positive for one person but negative for another. For example, in the Wilson, Kraft, and Dunn (1988a) study on political attitudes, what was easiest to verbalize or most available in memory about a particular political candidate may have been positive for some people (“He came across well on TV last week”) but negative for others (“He sleeps around”). Even if the same event is available in memory for all subjects, different subjects might interpret it quite differently, leading to attitude change in opposite directions (e.g., “He showed courage and leadership in his decision to sell arms to the Iranians,” vs. “He sure bungled that one”).

V. How and When Does Thinking about Reasons Influence Behavior?

To recap our argument to this point, we suggest that thinking about reasons brings to mind a biased sample of cognitions about the attitude object, producing attitude change in the direction of this biased sample. As reviewed earlier, this change seems to be expressed on self-report measures of attitudes but not on behavioral measures, resulting in reduced attitude–behavior consistency. In this section, we explain why thinking about reasons often does not influence behavior, and we outline the conditions under which it will.

One possibility invokes Campbell’s (1963) differential threshold argument. Campbell argued that different types of attitudinal responses have different thresholds; that is, some require more effort or motivation to perform. From lowest to highest threshold, he argued, are (1) autonomic–muscular reactions (e.g., galvanic skin response—GSR), (2) verbal reports about feelings toward the stimulus, (3) verbal reports about one’s behavioral intentions, and (4) overt, locomotor behavior. Some attitudes may be sufficiently strong to influence verbal reports, but not strong enough to be expressed behaviorally. Campbell (1963) reinterpreted some classic demonstrations of attitude–behavior inconsistency, such as LaPiere’s (1934), in terms of his differential threshold argument. Rather than being signs of logical inconsistency, he argued, some studies simply show that an attitude is weak, and thus it exceeds the threshold for verbal reports but not for behavior.

It is possible that the attitude change in our studies is another example of this phenomenon. Thinking about reasons may only cause weak, momentary attitude change, which is sufficient to change people’s attitude reports but is insufficient to budge them behaviorally. Such a view would be relatively trivial, demonstrating that we succeeded only in moving people a bit on 7-point attitude scales, but not on more consequential, behavioral dependent measures. There is reason to believe, however, that the differential threshold view cannot account fully for the effects of thinking about reasons on attitude–behavior consistency. When people have thought about reasons, the cognitive component of the attitude is at the forefront of their minds. If they then have the opportunity to behave toward the attitude object, we argue that this cognitive component will drive their behavior, at least initially. Why, then, have we found so much attitude–behavior inconsistency among people who have thought about reasons?

Our argument rests on our earlier suggestion that thinking about reasons turns attitudes into responses that are based primarily on cognitions about the attitude object. As argued by Millar and Tesser (1986), behavior as well as attitudes can be the result of either the affective or the cognitive component of an attitude. As seen shortly, Millar and Tesser (1986) found that thinking about reasons influenced behavior as well as attitudes, as long as the behavior was a function of the cognitions made salient by the reasons analysis manipulation. We agree that cognitively based behaviors can be influenced by thinking about reasons.
We suggest, however, that behavior that immediately follows a reasons analysis is apt to be cognitively driven, but that this behavior becomes more affectively driven over time. When people think about reasons, the cognitions that become salient may determine people’s attitude reports and their initial behavior. As they continue to interact with the attitude object, however, their basic affective reaction to it is more likely to reassert itself, hereby driving their behavior in a way that is likely to be inconsistent with their previously stated, cognitively based attitude.

An example of this process can be found in our studies that used Fazio’s puzzle paradigm. In these studies, subjects think about why they feel the way they do about five different types of puzzles. As a result of this reasons analysis, they use different criteria to evaluate the puzzles, and indeed end up evaluating them differently than do people in control conditions who do not think about reasons (Wilson & Dunn, 1986a; Wilson et al., 1984). After evaluating the puzzles, subjects are left alone during a free-time period, where they can play with any of the puzzles they choose.

We originally reported that subjects in the reasons-analysis condition choose to play with the same puzzles as subjects in the control conditions, and as a result, they exhibited less consistency between their evaluations and their behavior (Wilson & Dunn, 1986a; Wilson et al., 1984; see also Millar & Tesser, 1986). This is true if one averages subjects’ behavior toward the puzzles during the entire 15-minute free-time period. According to our new hypothesis about the sequence of behavior, however, we might expect that something like the following occurred: Subjects who thought about reasons brought to mind thoughts about the objective attributes of the puzzles, and they used these thoughts to evaluate them. When faced with the choice of which puzzles to play with in the free-time period, their initial choice was likely to be determined by these salient cognitions about the puzzles, and thus be consistent with their evaluations. As they played with these puzzles, however, their affect had a chance to reassert itself, and they discovered that they did not like these puzzles at all much after all (e.g., “This puzzle is challenging and is cleverly constructed, but you know, I think I really prefer some of the other ones”). Thus, their later choices of puzzles were more affectively driven, and thus inconsistent with their previous evaluations.

To test this possibility, we reanalyzed the data from the Wilson et al. (1984) and Wilson and Dunn (1986a) puzzle studies. First, we looked at subjects’ first choice of which puzzle to work on in the free-time period. As seen in Table III, subjects in the reasons conditions did tend to make different choices as to which puzzle to play with first than did subjects in their corresponding control conditions. This different pattern of first choices was not significant in the Wilson et al. (1984) study, which had a relatively small sample size, but it was significant in the Wilson and Dunn (1986a) study, which had a larger sample size. Combining across studies, using the method of adding z’s (Rosenthal, 1978), the difference was significant ($z = 2.25, p = .02$).

The differences reported in Table III do not indicate whether subjects’ first choices of puzzles corresponded to their liking ratings, as we have predicted. To test this hypothesis, we examined subjects’ reported liking for the first puzzle they chose to play with in the free-time period. If we are right—that subjects initially played with the puzzles they thought were the most interesting in both the control and the reasons conditions—there should be no difference in reported liking for the first puzzle they played with. This hypothesis was confirmed. The mean liking ratings of the first puzzles subjects played with, in the reasons versus the control conditions, was 5.46 versus 4.69 in the Wilson et al. (1984) study and 5.67 versus 5.72 in the Wilson and Dunn (1986a) study ($t’s < 1.17$).

We also expect that as subjects in the reasons condition played with the puzzles, their affective reactions reasserted themselves, causing them to end up playing with ones they had not rated very highly to begin with (accounting for the low attitude–behavior correlations we found in this condition). To test this possibility, we examined the ratings subjects gave to the puzzles they played with after their first choice, with the prediction that these ratings would be lower in the reasons than in the control conditions. This prediction was confirmed. The mean liking for the puzzles that subjects played with after their first choice, in the reasons versus the control conditions, were 3.28 versus 4.99 in the Wilson et al. (1984) study and 4.28 versus 4.50 in the Wilson & Dunn (1986a) study. Though the difference between conditions was significant only in the Wilson et al. (1984) study ($t(21) = 2.65, p < .02$), the overall difference combined across studies is significant ($z = 2.48, p < .02$).

Thus, we have found evidence for behavioral change as a result of thinking about reasons, suggesting that Campbell’s (1963) differential threshold view is not a satisfactory account of our results. An advantage of our hypothesis concerning initial versus later behavior change is that it specifies when behavior change should result from thinking about reasons and when it should not. If subjects have a relatively long period of time in which to interact with the attitude

<table>
<thead>
<tr>
<th>TABLE III</th>
</tr>
</thead>
<tbody>
<tr>
<td>PERCENTAGE OF PUZZLES PLAYED WITH FIRST, BY CONDITION</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>A</td>
</tr>
<tr>
<td>B</td>
</tr>
<tr>
<td>C</td>
</tr>
<tr>
<td>D</td>
</tr>
<tr>
<td>E</td>
</tr>
</tbody>
</table>
object, as in the puzzle studies and the Wilson et al. (1984) couples study (in which the behavioral measure was whether subjects had broken up several months later), then people’s affective response toward the stimulus has a chance to reassert itself, resulting in attitude–behavior inconsistency. If the behavioral measure is a short-term one, such as subjects’ initial choice between different attitude objects, this choice might well be driven by the new, cognitively based attitude, resulting in attitude–behavior consistency.

Such a result was found in a recent study by Wilson, Lisle, and Schooler (1988c). Subjects in this study evaluated a set of five posters of the sort commonly found in college students’ rooms (the attitude measure), and they were allowed to choose one to take home (the behavioral measure). Because the behavioral measure was not a long-term one, in which affect had a chance to reassert itself, we predicted that thinking about reasons would influence both subjects’ liking ratings and their choice of which poster to keep. This prediction was confirmed. Interestingly, as is discussed in detail in a later section, subjects in the reasons condition came to regret their choice of poster more after taking it home. As our time-course hypothesis suggests, subjects’ affect toward the posters had more of a chance to reassert itself once they got the posters home, thus they were more apt to regret a choice that had been cognitively determined.

This hypothesis can be used to reinterpret some of the results of our earlier studies summarized in Table 1. As already seen, one reason the Fazio puzzle paradigm has been so successful in demonstrating that thinking about reasons reduces attitude–behavior consistency is that the behavioral measure was a long-term one. Further, it may explain why the results of Study 1 by Wilson and Dunn (1986a) were weak. In this study, subjects waiting in line at a college dining hall were asked to think about why they liked or disliked several beverages, and the behavioral measure was which beverages they purchased with their dinners. Because the behavioral measure was a short-term one, we might expect that it would be influenced by any attitude change that occurred as a result of thinking about reasons. As predicted, attitude–behavior consistency was relatively high in the reasons conditions of this study.

Some of the other studies in Table 1 are difficult to classify in terms of our hypothesis about the time frame of the behavior. These studies used facial expressions while watching a slide show as the behavioral measure, which seemed not to have changed as a result of thinking about reasons. Unlike the other studies, however, the behavior measure was assessed before the self-report measure, preventing a test of the hypothesis that behavior was initially consistent with self-reports but became less so as time passed. Another study reported in Table 1 seems inconsistent with the time-frame hypothesis: Study 2 by Wilson, Kraft, and Dunn (1988a). Here, the behavioral measure was the number of fliers subjects took for the 1984 Presidential candidates. This appears to be a short-term behavioral measure, yet the number of fliers subjects took in the reasons condition was not consistent with their reported attitude. The reason for this may be due to Miller and Tesser’s (1986) distinction between behaviors that are affectively versus cognitively driven. Some behaviors, such as seeing a picture of Mondale and Ferraro and having to commit oneself to posting these pictures around the campus, may elicit an immediate affective response that is inconsistent with one’s cognitions about the attitude object.

To summarize, we suggest that when people are asked to think about reasons, they feel compelled to give a good story as an explanation of their feelings. These reasons are often a biased sample of cognitions about the attitude object, which imply a different attitude than subjects held previously. Unaware that their reasons are a biased sample, subjects change their attitude in the direction implied by their reasons. The attitude change that occurs is not necessarily in a uniformly positive or negative direction. Wilson, Kraft, and Dunn (1988) and Wilson and Kraft (1988) found that some subjects changed in a positive direction, others in a negative direction as a result of thinking about reasons. The change that occurs seems to direct subjects’ initial behavior, as indicated by our reanalysis of the Wilson et al. (1984) and Wilson and Dunn (1986a) puzzle studies. Overall, however, behavior seems to snap back, possibly because subjects’ affective response toward the stimulus has a chance to reassert itself.

One advantage of this account is that it implies several boundary conditions on the disruptive effects of thinking about reasons. Before discussing these, however, it is important to discuss an alternative explanation of the effects of thinking about reasons, which Millar and Tesser (1986) have offered.

MILLAR AND TESSER’S (1986) MISMATCH HYPOTHESIS

Millar and Tesser (1986) have recently advanced a somewhat different account of why thinking about reasons reduces attitude–behavior consistency. Consistent with our arguments, they suggest that analyzing reasons causes people to think about attributes of the attitude object, which increases the salience of the cognitive component of the attitude. Also similar to our position, they argue that people’s attitude is a function of those thoughts or feelings about the attitude object that are salient in memory at any given time. Their focus, however, is on the traditional breakdown of attitudes into affective and cognitive components. Rather than saying that attitude change results from bringing to mind a biased sample of cognitions, they suggest that thinking about reasons increases the salience of the cognitive component of an attitude at the expense of the affective component.

More specifically, Millar and Tesser (1986) suggest that (1) the attitudes people report depend on which attitude component is salient in memory, because the affective and cognitive components of attitudes often conflict; (2) thinking about reasons increases the salience of the cognitive component, such that attitude reports become cognitively driven; (3) behavior is often affectively driven (as was the case, they argue, in previous reasons analysis studies); (4) when people think about reasons, attitude–behavior consistency will thus suffer, because attitude reports become a function of the cognitive component of attitudes, whereas behavior is a function of the affective component. This reasoning explains why
Wilson and Dunn (1986a) found different effects of analyzing reasons versus focusing on affect. The attitude reports of people who thought about reasons became cognitively driven, whereas their behavior remained affectively driven, resulting in low attitude–behavior correlations. When people focused on their affect, the affective component of their attitude determined their attitude reports, which thus matched their affectively driven behavior.

The mismatch hypothesis received additional support in an experiment by Millar and Tesser (1986). As in the Wilson and Dunn (1986a) study, some subjects thought about why they liked or disliked a set of puzzles, while others focused on how they felt about the puzzles, without justifying their feelings. The reasons-analysis manipulation increased the salience of the cognitive component of subjects’ attitudes, reasoned Millar and Tesser (1986), while the affective-focus manipulation increased the salience of the affective component. In addition, Millar and Tesser manipulated whether subjects’ behavior toward the puzzles was cognitively or affectively driven. In the affective-behavior condition, subjects were allowed to work on any of the puzzles they chose during a free-play period (as in Wilson and Dunn’s study), with the assumption that their choice of puzzles was motivated chiefly by a desire to play with the ones they found most pleasing. In the cognitive-behavior condition, subjects were told that the puzzles were designed to increase their analytic ability, and that they would receive a test of analytic ability immediately after the free-time period. The assumption here was that the subjects’ choice of puzzles would be a function of their beliefs about which puzzles would best prepare them for the analytic ability test, and thus would be more cognitively driven.

Millar and Tesser’s (1986) 2 × 2 design created two cells in which the components driving attitudes and behavior matched (one in which both attitudes and behavior were cognitively driven, and one in which both attitudes and behavior were affectively driven), and two mismatch cells, in which attitudes and behavior were each driven by different components. The main dependent measure was the mean within-subject correlation between each subject’s ratings of the puzzles and the amount of time he or she spent playing with each one in the free-time period. As predicted, and as seen in Fig. 1, attitude–behavior consistency was high in the matching cells and low in the mismatch cells. It thus appears that another way in which thinking about reasons lowers attitude–behavior consistency is that people phrase their attitude reports in cognitive terms, while (often, at least) their behavior is a function of their affect.

We view the Millar and Tesser position and ours as complementary, with empirical evidence supporting both. The chief difference is as follows: While both positions argue that attitudes become more cognitively based as a result of thinking about reasons, we suggest: that this is because people bring to mind a biased sample of cognitions, while Millar and Tesser (1986) argue that a preexisting cognitive component is emphasized at the expense of a conflicting affective component. This is potentially an important distinction, because it implies different boundary conditions on the effects of thinking about reasons. According to our position, thinking about reasons will be particularly disruptive when an attitude is primarily affective in nature, with few supporting cognitions. Under these conditions, we suggest, people are not used to phrasing their attitude in cognitive terms, and forcing them to do so is most likely to be disruptive—that is, they are most likely to generate a set of cognitions that is a biased sample of the factors underlying their attitudes. Evidence for this hypothesis is presented shortly. According to the Millar and Tesser position, thinking about reasons should be most disruptive when people initially have conflicting affective and cognitive components to their attitudes. A subsequent set of studies by Millar and Tesser (1988) found support for this hypothesis. Thus, our position and Millar and Tesser’s appear to be nonoverlapping and noncontradictory, with empirical support for both.

A second difference between our position and Millar and Tesser’s (1986) concerns the conditions under which the cognitively based attitude that results from thinking about reasons will conflict with behavior. Millar and Tesser (1986) suggest that behaviors that are affectively determined will conflict with this new attitude, whereas behaviors that are cognitively driven will not. We argue that, in addition to the nature of the behavior (whether it is affectively or cognitively determined), one must consider its time course. People’s initial behavioral response, we suggest, will often follow from their new, cognitively based attitude, whereas over time, behavior becomes more affectively driven.

Fig. 1. Effects on attitude–behavior consistency of (1) analyzing reasons versus focusing on affect and (2) cognitively versus affectively mediated behavior. Adapted from Millar and Tesser with permission. Copyright 1986 by American Psychological Association.
VI. Boundary Conditions on the Effects of Thinking about Reasons

Any good explanation of an effect specifies the boundary conditions that limit it. Several such boundary conditions are implied by our explanation of why thinking about reasons can change attitudes and lower attitude-behavior consistency, and there is empirical evidence for some of these delimiting conditions.

A. AFFECTIVELY VERSUS COGNITIVELY BASED ATTITUDES

We suggested earlier that some types of attitudes may be more susceptible to the effects of thinking about reasons than others. Any attitude that people find easy to explain, so that they do not come up with a biased set of cognitions, should be relatively unaffected by thinking about reasons. An example of such an attitude, we suggest, is one that is based on a relatively small number of well-thought-out cognitions about the attitude object, such that people can easily call upon these cognitions when asked to explain the attitude. Other kinds of attitudes are more difficult to explain. For example, some attitudes have a strong-amorphous affective component with few supporting cognitions, and it is difficult to know where this affect came from. Attempts to explain this type of attitude may be particularly apt to lead to a biased set of reasons. Most of the attitudes examined in our studies seem to be of this latter type, such as people’s feelings about their dating partner or their liking for a set of vacation pictures.

This distinction is similar to one made by Zajonc (1980), Zajonc and Markus (1982), and Zanna and Rempel (in press) concerning the affective and cognitive underpinnings of attitudes. These authors argue, in contrast to the traditional attitude theory, that some attitudes are primarily affectively based, in that they have strong affective components without a large number of beliefs supporting the attitude. As discussed earlier, attitudes can be affective because they are the result of minimal cognitive processing, such as mere exposure, classical conditioning, or operant conditioning, or because they stem from people’s core values such as their moral or religious beliefs. We refer to any attitude that is grounded by strong emotions rather than a logical analysis of the evidence as “affectively based.”

Other attitudes conform more to the traditional definition, in that they have strong affective and cognitive components. (Whether an attitude can consist of only a cognitive component—with little or no accompanying affect—is an interesting question that is not considered here.) Many of these, which we refer to as “cognitively based attitudes,” result from a careful analysis of the relevant facts. For example, people might evaluate a political candidate by compiling and evaluating information about the candidate’s stance on the issues. Such attitudes are typically accompanied by affect, but the affect follows from a careful analysis of the attributes of the attitude object, and it is likely to be subordinate to the cognitive component of the attitude.

Earlier, we argued that thinking about reasons is disruptive because people construct an explanation of their feelings, and this explanation often implies a somewhat different evaluation of the attitude object than they previously held. This process would only be expected to occur, however, if people were not fully aware of what had caused their affect, forcing them to construct the most plausible explanation. If people knew why they felt the way they did, then the reasons they brought to mind would be the correct ones, and the reasons would be consistent with their affect. We suggest that people with cognitively based attitudes can more easily recall the correct reasons for their attitudes because they are more likely to have worked out carefully how they feel by deliberating about the attributes of the attitude object. Thus, they should be less susceptible to the disruptive effects of analyzing reasons.

People are less likely to know the determinants of their affective attitudes. Important determinants such as mere exposure are apt to be overlooked because they are implausible; that is, they are not part of their implicit theories about what causes people to feel the way they do (Nisbett & Wilson, 1977b). For example, few people recognize the extent to which their political preferences result from the party identification of their parents (Campbell, Converse, Miller, & Stokes, 1960; Hyman, 1959). Even if people originally knew the determinants of their feelings, as time passes, it is often the affect that remains in memory while its causes are forgotten (Lingle & Ostrom, 1981; Zajonc, 1980). Further, if the processes underlying an affective reaction are minimal or inaccessible (as in the mere exposure effect), it is unlikely that people could easily verbalize the nature of these processes. Finally, attitude objects are often complex and multidimensional, and it is simply too difficult to sort out from among the many possibilities the extent to which different attributes influence our feelings, as in our earlier example of people’s inability to know exactly why they feel the way they do about their spouses.

Thus, people are more apt to be misled when trying to explain affectively based attitudes, increasing the likelihood that they will come up with a biased set of reasons. Further, because affective attitudes are dominated by a strong, amorphous affective component, the very act of trying to translate this affect into a logical-sounding verbal code may be disruptive. Attempting to recast such attitudes in cognitive terms probably contributes to the distortions and disruptions that occur. In short, we argue that thinking about reasons can turn a gut reaction into an opinion based on a biased sample of cognitions about the attitude object.

When people already have an informed opinion (i.e., a cognitively based attitude that is already phrased in a logical, verbal code), having to verbalize reasons should be easier and less disruptive.
We do not mean to argue that people are unaware of the causes of all of their affectively based attitudes and fully aware of the causes of all of their cognitively based attitudes. Sometimes people have an affective reaction resulting from a single, powerful source, and this source is easy to identify. Zanna and Rempel (in press), for example, present the case of the attitude toward drunk driving of a parent whose child was killed by an intoxicated driver. The parent’s attitude is likely to be based on an intense emotional reaction, not on an appraisal of facts and statistics about accident rates of intoxicated drivers; nonetheless, the parent will know full well the reason for the attitude. Conversely, there are undoubtedly cognitively based attitudes that people have difficulty explaining, because they consist of so many complex beliefs that it is difficult to know how much each one contributes to the attitude. Our claim is that people are more likely to know why they feel the way they do if their attitude is based on a careful appraisal of the attributes of the attitude object than if it is not; that is, the dimension of affectively versus cognitively based attitudes is correlated with people’s ability to explain them accurately, though not perfectly so.

A thorny issue with regard to the distinction between affectively and cognitively based attitudes is whether they differ only in terms of the nature of their underpinnings, or whether they also differ in terms of their strength. Are affectively based attitudes typically weaker or stronger than cognitively based ones? Are they more or less accessible or available in memory? These are important questions, because how strong or accessible an attitude is may also determine its susceptibility to disruption by thinking about reasons.

B. WEAK VERSUS STRONG ATTITUDES?

Just because an attitude is not based on a set of well-thought-out cognitions does not mean that it is affectively based. There are many issues about which people care little and do not have well-articulated attitudes. This type of response is unlikely to have strong cognitions or strong affect, and it is best characterized as a nonattitude (Converse, 1970; Hovland, 1959) or a weak attitude with low accessibility (Fazio, 1986; Fazio & Zanna, 1981). Thus, there are at least three types of attitudes: (1) those that are strongly held and based on a set of well-thought-out cognitions about the attitude object (cognitively based attitudes), (2) those that are strong but grounded primarily on an affective reaction to the stimulus (affectively based attitudes), and (3) those that are weak and inaccessible, with few supporting cognitions or affective bonds (nonattitudes).

It is possible that attitudes that are weakly held are also highly susceptible to a reasons-analysis manipulation. This argument is based on Fazio’s model of attitude accessibility (Fazio, 1986). Fazio distinguishes between attitudes that are based on direct experience with the attitude object (and thus are strong and accessible) versus attitudes that are not based on direct experience (and thus are weak and inaccessible). The latter type of attitude has been demonstrated (1) to be less likely to predict behavior, (2) to be held with less confidence, and (3) to take longer to be expressed than strong attitudes (Fazio, 1986; Fazio & Zanna, 1981): Generalizing from Fazio’s work, weak attitudes may also be more susceptible to alteration by attitude-change techniques, including our reasons-analysis manipulation.

It may well be the case that weak attitudes are more susceptible to change. Our studies suggest, however, that our effects are not limited to weak attitudes; the effects can occur with relatively strong attitudes. In virtually all of our studies, we began with attitudes that predicted behavior well (and thus were strong, according to the Fazio model), and we found that thinking about reasons undermined attitude–behavior consistency. Indeed, many of these studies used a paradigm developed by Fazio to create strong attitudes, in which subjects were given direct experience with a set of puzzles. Thinking about reasons has been found to change people’s attitudes toward these puzzles and to reduce the consistency between these new attitudes and their subsequent behavior, even though the attitudes were strong at the outset.

Instead of weak attitudes being those that are most susceptible to the disruptive effects of thinking about reasons, we suggest that it is those that are affectively based. It is not a simple matter, however, to operationalize the difference between an attitude that is weak versus one that is strong but affectively based. Measures have been developed to assess separately the affective and cognitive components of attitudes (e.g., Rosenberg, 1956, 1960), but responses on these measures might not reflect the extent to which an attitude is based on one or the other component. In principle, people could know a good deal about the issues in a political campaign, for example, but their attitudes could still be based on intense emotional reactions to the candidates rather than on their beliefs about the issues.

Indirect evidence for the existence of such attitudes does exist, chiefly from studies that have shown that it is the affective component of an attitude that correlates the best with people’s behavior and preferences (Abelson, Kinder, Peters, & Fiske, 1982; Seligman et al., 1979). It is not an easy matter to demonstrate, however, that a particular attitude is affectively based, and it is particularly difficult to demonstrate that two attitudes are equal in strength but differ in the extent to which their underpinnings are affective.

We return to a discussion of this difficulty after describing several tests of the hypothesis that thinking about reasons will only influence attitudes that are affectively based. Our first four tests divided subjects on the basis of their knowledge about or experience with the attitude object. We assumed that people who were unknowledgeable or inexperienced would be more likely to have affectively based attitudes. There is an obvious danger to operationalizing affectively versus cognitively based attitudes by dividing people on the basis of their experience with the attitude object, because people’s attitudes may differ in many respects other than their affective basis, such as their strength or accessibility. Most of the attitude objects we chose, however, were of sufficient importance and salience to people that it seemed unlikely that people would have weak attitudes. Some
evidence for this assumption is presented, as is a fifth study that manipulated whether attitudes were affectively or cognitively based.

C. EMPIRICAL TESTS

1. The Couples Study Revisited

People who have a lot of experience with an attitude object are more likely to have opinions that are based on a careful appraisal of the attributes of the attitude object, thus any introspection that brings these attributes to mind should have little impact. People who are inexperienced should be more likely to have affectively based attitudes that are the result of quick appraisals or such factors as classical conditioning or mere exposure. These people probably are less knowledgeable about why they feel the way they do, and they are more apt to be misled by being asked to give reasons.

For example, the participants in the Wilson et al. (1984) couples study differed in the amount of time they had been dating. In a reanalysis of this study, we performed a median split on the length of the couple's relationship (median = 5.1 months), with the assumption that those who had been dating for a relatively short time would be more likely to have affectively based attitudes. It is possible that instead they had weak attitudes. Anyone who has observed an 18-year-old in love would probably agree, however, that these attitudes can be very strong, but they may consist of relatively few cognitions. The couples who had been dating for a relatively long time were assumed to have attitudes that were more cognitively based, in the sense that they had had more opportunity to develop a set of cognitions about their partners, and they could more easily verbalize these thoughts.

We predicted that our reanalysis of the Wilson et al. (1984) study would show that thinking about reasons reduced attitude–behavior consistency only for those couples who had been dating for a relatively short period of time. As seen in Table IV, this prediction was confirmed. An a priori contrast was performed on the correlations in Table IV that assigned a weight of −3 to the reasons–new relationship group and a weight of 1 to the other three groups. This contrast was significant when individuals were the unit of analysis and marginally significant when couples were the unit of analysis (z-s = 2.55 and 1.74, p's = .01 and .08, respectively).

2. Attitudes toward Walter Mondale

Wilson, Kraft, and Dunn (1988a) conducted a series of additional tests of our hypothesis that thinking about reasons is most apt to influence affectively based attitudes. In one study, subjects were divided on the basis of their knowledge about the candidates in the 1984 Presidential election. Half of the subjects in each group were asked to think about why they felt the way they did about Walter Mondale, while the others completed a filler questionnaire. The chief dependent measure was attitude–behavior consistency (the behavioral measure of subjects' attitude was the number of Mondale–Ferraro fliers they were willing to take to post around the campus). Thinking about reasons again decreased consistency for unknowledgeable subjects but not for knowledgeable ones (see the first row of Table V). A planned comparison that assigned a weight of −3 to the reasons–unknowledgeable cell and weights of 1 to the other three cells was significant (z = 3.41, p = .001).

3. Attitudes toward the 1988 Presidential Candidates

The two studies just reviewed showed that knowledge moderates the effects of thinking about reasons on attitude–behavior consistency. The next two studies tested the hypothesis that knowledge also moderates the effects of thinking about reasons on attitude change. As reviewed earlier, Wilson, Kraft, and Dunn (1988a) performed a study in which attitudes toward six of the candidates for President

<table>
<thead>
<tr>
<th>TABLE IV</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Effects of Analyzing Reasons and Length of Dating Relationship on Attitude–Behavior Consistency</strong></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Individuals as the unit of analysis</td>
</tr>
<tr>
<td>Couples as the unit of analysis</td>
</tr>
</tbody>
</table>

*Data from Wilson et al. (1984, Study 3). Couples were classified as having "newer" or "longer" relationships on the basis of a median split on the length of time they had been dating (median = 5.1 months).

<table>
<thead>
<tr>
<th>TABLE V</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Effects of Analyzing Reasons and Knowledge on Attitude–Behavior Consistency and Attitude Change</strong></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Attitude–behavior consistency</td>
</tr>
<tr>
<td>Absolute value of change in attitude</td>
</tr>
</tbody>
</table>

*Entries in the first row are from Wilson, Kraft, and Dunn (1988a, Study 1); entries in the second row are from Wilson, Kraft, and Dunn (1988a, Study 2).
in 1988 were measured twice, once in a mass testing session and once in the laboratory. In our earlier discussion of this study, we reported that subjects who thought about reasons were significantly more apt to change their minds about the candidates than subjects in the control condition. These results, however, were reported only for subjects who were relatively unknowledgeable about the candidates. As predicted, there was not a significant difference between the absolute value of change among those subjects who were knowledgeable about the candidates, as seen in the second row of Table V. An analysis of variance on these change scores yielded a significant Reasons Condition X Knowledge interaction (p < .05), though our standard planned comparison, as described previously, was significant only at the .12 level.

4. Attitudes toward Posters

The Wilson, Lisle, and Schooler (1988c) study allows a further test of the hypothesis that knowledge about the attitude object moderates the effects of thinking about reasons on attitude change. In brief, subjects were asked to examine and rate their liking for five posters that were of two different types: Two were reproductions of Impressionist art, whereas three were of a more "pop" style, such as a photograph of a cat with the caption, "Gimme a Break." We predicted that under normal conditions, the art posters would be the most affectionately pleasing, and thus most preferred by subjects. When subjects thought about why they liked or disliked the posters, however, we predicted that they would focus on their cognitions about the posters, which would be easier to verbalize about the popular ones. Thus, we expected that subjects in the reasons-condition would change their attitudes toward the posters, becoming more positive toward the popular ones and less positive toward the art posters.

This prediction was made, however, only for people who were relatively unknowledgeable about art, and thus should find it most difficult to verbalize their cognitions. People who are knowledgeable about art should find it easier to verbalize why they feel the way they do, and thus should be less apt to bring to mind a biased set of reasons. This prediction was confirmed. For unknowledgeable subjects, thinking about reasons caused an increase in liking for the popular posters and a decrease in liking for the art posters. Thinking about reasons had relatively little impact on knowledgeable subjects. A contrast that tested the overall predicted pattern of results was highly significant (F[1,31] = 12.96, p = .001).

Thus, four studies have found that knowledge about the attitude object moderates the effects of thinking about reasons. The explanation for these findings, we suggest, is that unknowledgeable people had attitudes that were more affective in nature, and thus more apt to be disrupted by an attempt to explain them. Clearly, however, people who are unknowledgeable might differ in other ways from people who are knowledgeable. It would thus be desirable to manipulate the basis of an attitude, rather than relying on an individual difference variable. This strategy was used in the next study reviewed here.

5. Manipulating the Basis of an Attitude

A study by Wilson and Kraft (1988) attempted to create attitudes toward the same stimuli that were affectively based for some people and more cognitively based for others. Subjects were told that the purpose of the study was to look at the kinds of impressions people form of clinical psychology graduate students, to see if these impressions were related to the students' therapeutic skills. As a means of creating affectively based attitudes toward the students, some subjects were exposed to a mere exposure manipulation, in which they were given no information about the students other than their photographs. The frequency with which subjects saw these photographs was manipulated, so that subjects saw the photographs of some students only once, others 25 times. Subjects in the cognitive-information group received brief profiles about each student that described her career goals, research interests, and leisure-time activities, in the absence of any photographs. These profiles were pretested such that they caused a range in liking for the students equal to the range in the mere exposure condition.

Half of the subjects in each condition received our standard reasons manipulation, in which they wrote down why they liked or disliked each of the clinical psychology graduate students. Subjects in the control condition completed a filler questionnaire. All subjects then rated their liking for the students on 7-point scales. We predicted that the reasons manipulation would have relatively little impact on the subjects in the cognitive-information condition, because they would have formed their attitudes by integrating a relatively small number of pieces of information about the students (e., "She likes to bowl and read science fiction books"). When asked to give reasons, it should have been relatively easy for these people to verbalize why they felt the way they did.

It should have been more difficult for subjects in the mere exposure condition to verbalize why they felt the way they did, for at least two reasons: (1) It is unlikely that subjects knew that their attitudes were influenced by the frequency with which they saw the photographs, thus they were unaware of the major determinant of their attitudes; (2) the nonverbal information they received (i.e., the photographs) was more difficult to verbalize than were the profiles. Therefore, subjects in the mere exposure condition were hypothesized to be more apt to be misled by thinking about reasons, and thus more apt to change their minds about how they felt.

This prediction was confirmed, as seen in Table VI. For subjects who saw the profiles, thinking about reasons had little effect on their attitudes (p = .37). For subjects in the mere exposure condition, however, thinking about reasons changed their attitudes in a negative direction (p < .001): A contrast testing the overall prediction that thinking about reasons would not affect attitudes in the
basis of attitudes, and we found that those resulting from mere exposure were changed by thinking about reasons, whereas those resulting from reading the profiles were not. Finally, in some of our earlier studies, we had created strong attitudes by giving subjects direct experience with the attitude object, using Fazio’s procedure for doing so (e.g., Regan & Fazio, 1977). Thinking about reasons was found to change attitudes in these studies and to undermine attitude–behavior consistency, suggesting that it is not just weak attitudes that are susceptible to these effects. On balance, then, the evidence points to our distinction between affectively and cognitively based attitudes as a moderator of the effects of thinking about reasons. This issue is far from settled, however. It is not a simple matter just to assess the affective and cognitive bases of attitudes, and to distinguish these characteristics from attitudinal strength. A fertile area for future research would be to develop such assessment techniques.

Our research indicating that some types of attitudes are more likely to be changed by thinking about reasons, it should be noted, helps explain the inconsistency we referred to earlier between our self-persuasion hypothesis and consistency theories. We have argued that when people are asked to explain an attitude, they often come up with a biased set of cognitions about the attitude object, which causes them to change their minds about how they feel. As noted earlier, consistency theories suggest that when people are asked to explain an attitude, they would think of reasons that are consistent with the phenomenon they are attempting to explain.

People will generate a consistent set of beliefs, we argue, for some types of attitudes—namely those that are cognitively based—and thus are already grounded on a set of well-thought-out, integrated beliefs. When people are asked to explain an affectively based attitude, however, they do not have a ready-made set of consistent beliefs to call upon. In fact, there is evidence that when people are relatively unknowledgeable about an attitude object—and thus are likely to have affectively based attitudes—their beliefs about the attitude object are relatively inconsistent (Lusk & Judd, 1988). As we have seen, the reasons people do come up with are also likely to be biased in a number of respects. Previous studies have shown that if people are duped, through linguistic devices, into generating a set of biased reasons, they adopt the attitude implied by these reasons (Salancik, 1974; Seligman et al., 1980). It is thus not surprising that if people come up with a biased sample of reasons on their own, a similar type of attitude change occurs.

D. PREDICTING BEHAVIOR THAT IS AFFECTIVELY VERSUS COGNITIVELY DRIVEN

An additional boundary condition was suggested by Millar and Tesser (1986). They demonstrated that thinking about reasons reduces the ability of attitudes to
predict affectively based behavior but not cognitively based behavior. Because analyzing reasons heightens the salience of the cognitive component of the attitude, behavior that is motivated chiefly by these cognitions will follow from attitude reports. Millar and Tesser (1986) did not have a no-introspection control group, thus it is unclear whether thinking about reasons actually increased correlations with cognitively based behaviors or simply did not lower them. In either case, an important boundary condition has been established.

It will take further research to delineate better which types of behavior are motivated primarily by affect and which by the cognitive component of attitudes. Millar and Tesser (1986) describe affectively based behaviors as “consummatory” (i.e., those performed to obtain pleasure or avoid pain) and cognitively based behaviors as “instrumental” (i.e., those performed to achieve some goal associated with the attributes of the attitude object, such as working on puzzles to develop one’s analytic ability). As with the basis of attitudes, it would be extremely difficult to estimate what percentage of people’s behavior is motivated chiefly by affect versus cognition. Enough behavior seems to emanate from the heart rather than the head, however, that this boundary condition does not trivialize the effects of thinking about reasons.

E. OTHER BOUNDARY CONDITIONS

There are several other potential boundary conditions of thinking about reasons that have yet to be tested empirically. For example, the reasons analyses performed by subjects in our studies were relatively brief. Perhaps if people were to spend more time analyzing their motives, and if they were more motivated to do so (e.g., if they were facing an important decision), they could more successfully integrate with their affect the cognitions that result from thinking about reasons. Similarly, the analysis of reasons in our studies has been carried out solitarily. It may be that if we discuss our motives with another—such as a friend, spouse, or psychotherapist—the other person can be more objective, seeing inconsistencies and distortions that are not obvious to us. Finally, it may be that some people are better at introspection than others. Just as there are individual differences in the extent to which people focus on themselves (Fenigstein, Scheier, & Buss, 1975), there might be differences in skill at analyzing reasons. We included Fenigstein et al.’s (1975) private self-consciousness scale in one of our reasons-analysis studies (Wilson et al., 1984, Study 1), and we found that thinking about reasons lowered attitude–behavior correlations both for people high and for people low in private self-consciousness. A more direct measure of skill at introspection (rather than a measure of the frequency of self-focus), however, might predict which individuals are most susceptible to the disruptive effects of thinking about reasons.

VII. Alternative Explanations for the Disruptive Effects of Thinking about Reasons

There are other possible explanations for the disruptive effects of thinking about reasons. Four such possibilities are considered here, along with evidence for and against each one.

A. DOES THINKING ABOUT REASONS INCREASE DISCREPANCIES BETWEEN THE AFFECTIVE AND THE COGNITIVE COMPONENTS OF ATTITUDES?

We argued earlier that thinking about reasons causes people to recast their attitudes in cognitive terms, and that these attitudes then conflict with behavior that is affectively based (particularly when enough time has elapsed for people’s affective reaction to reassert itself). We have also tested an additional hypothesis about why thinking about reasons can lower attitude–behavior consistency. When people have a cognitively based attitude, we suggest, they are more apt to have a set of cognitions about the attitude object that are well integrated with their affect. For example, people whose attitudes toward a political candidate are based on a careful appraisal of the candidate’s stance on the issues are likely to have cognitions that are consistent with their affective reaction, because their affect stems from these cognitions.

When people have affectively based attitudes, they are less likely to have spent time integrating their thoughts with their feelings, and they are thus more likely to have beliefs that conflict with their affect (e.g., “He may have voted the wrong way on a few bills in Congress, but I really like him”). Under normal conditions, people with affectively based attitudes probably focus on those beliefs that are consistent with their affect; that is, the affect swamps the cognitions. When asked to explain their attitudes, however, they may focus on their beliefs, highlighting their inconsistency with their affect (“Why do I feel the way I do? Well, now that I think about it, his position on some key issues is different from mine”). If so, thinking about reasons may not only change people’s minds about how they feel, it may also create an attitude with discrepant affective and cognitive components. This would help explain why thinking about reasons reduces attitude–behavior consistency, because there is ample evidence that attitudes with discrepant components are unstable and poor predictors of behavior (Norman, 1975; Rosenberg, 1968).

Evidence for part of this proposed sequence of events was recently obtained by Lusk and Judd (1988). In two studies, they found that subjects who were relatively knowledgeable about politics had beliefs about political candidates that
were intercorrelated; that is, their knowledge about the candidates was integrated into a schema that was evaluatively consistent. Subjects who were unknowledgeable had beliefs that were more apt to be inconsistent. To the extent that the unknowledgeable people had attitudes that were more affectively based, this supports our contention that affectively based attitudes are less apt to be supported by a set of consistent beliefs.

To see if thinking about reasons highlights discrepancies between the affective and cognitive components of affectively based attitudes, we measured the affective and cognitive components of attitudes separately in four of our studies: a survey of attitudes toward marijuana by Wilson and Dunn (1986b); Wilson, Kraft, and Dunn’s (1988a) study of attitudes toward the 1984 and 1988 Presidential candidates; and Wilson Kraft’s (1988) study that manipulated the affective versus cognitive basis of attitudes. In the first three of these studies, subjects were divided on the basis of their knowledge about or experience with the attitude object, as a means of operationalizing the affective versus cognitive basis of their attitudes. In the fourth study, we manipulated the basis of their attitudes, as discussed previously.

In the last three studies, the affective and cognitive components of attitudes were assessed using techniques described by Rosenberg (1960). In the first, the affective component was assessed with semantic differential scales, and the cognitive component was assessed with a previously developed measure of beliefs about marijuana (Schlegel, 1975). The consistency between subjects’ affective and cognitive responses was assessed by standardizing their responses on each measure, and then computing the absolute value of the difference between them.

Thinking about reasons increased the discrepancy between the affective and cognitive components of attitudes in three of the four studies when attitudes were affectively based (see Table VII). Averaging across studies, this difference was significant ($z = 2.02, p = .04$). Thinking about reasons had relatively little impact on attitudes that were cognitively based. A contrast testing the overall predicted pattern of results (with a contrast weight of 3 in the reasons–affective attitude cell and –1 in each of the remaining cells) was also significant across studies ($z = 2.93, p = .003$). These differences, however, were not very large. Even though the difference between the affective–reasons and affective–control conditions was significant across studies, it failed to reach significance in any of the individual studies. Further, as seen at the bottom of Table VII, the difference in discrepancy scores between these conditions averaged only .10 standard deviation units.

Thus, some weak support for our hypothesis concerning the consistency between the affective and cognitive components has been obtained. Clearly this effect is not very strong, however, and it is unlikely that it fully accounts for the reductions in attitude–behavior consistency we have found. As discussed earlier, these reductions appear to be due in large part to the fact that when people think about reasons they recast their attitudes in cognitive terms, though their behavior is often affectively driven.

**B. DOES THINKING ABOUT REASONS LOWER CONFIDENCE IN ATTITUDES?**

Analyzing the reasons for one’s feelings might cause people to consider alternative points of view, shaking their confidence in their attitudes. Unable to explain fully why she is attracted to her dating partner, for example, a reason might begin to question the strength or nature of her feelings. Fazio and Zanna (1981) demonstrated that attitudes held with low confidence are poor predictors of behavior, thus a loss of confidence precipitated by thinking about reasons would explain the reductions in attitude–behavior consistency found in our studies.

This hypothesis is implied in the earlier quote from Mario Vargas Llosa, who “could no longer easily tell what I liked or didn’t” after analyzing the films at the Berlin Film Festival. It is echoed by a character in a novel by Iris Murdoch, who observed that “not everything is improved and clarified by being dug up” (Murdoch, 1986, p. 422). Finally, in the first study (of which we are aware) that used a reasons-analysis manipulation, some evidence was found for this hypothesis. Carper and Doob (1953) asked subjects for their level of agreement with six attitude statements having to do with the equal access to governmental services by American citizens. Half of the subjects were first asked to explain why they felt the way they did, whereas the others were not. Interestingly, those who thought about reasons were more apt to erase their subsequent attitude

<table>
<thead>
<tr>
<th>Table VII</th>
</tr>
</thead>
<tbody>
<tr>
<td>Effects of Thinking about Reasons and Knowledge on the Consistency between the Affective and Cognitive Components of Attitudes</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Affectively based</th>
<th>Cognitively based</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Control</td>
<td>Reasons</td>
</tr>
<tr>
<td>Wilson &amp; Dunn (1986b)</td>
<td>.48</td>
<td>.56</td>
</tr>
<tr>
<td>Wilson, Kraft, &amp; Dunn (1988a), Study 1</td>
<td>.63</td>
<td>.85</td>
</tr>
<tr>
<td>Wilson, Kraft, &amp; Dunn (1988a), Study 2</td>
<td>.70</td>
<td>.65</td>
</tr>
<tr>
<td>Wilson &amp; Kraft (1988)</td>
<td>.62</td>
<td>.79</td>
</tr>
<tr>
<td>Mean</td>
<td>.61</td>
<td>.71</td>
</tr>
</tbody>
</table>

*Cell entries are the absolute value of the difference between subjects' standardized affective response and standardized cognitive response.
responses than those who did not, suggesting that they became less certain about how they felt.3

We have tested the reduced-confidence hypothesis in several of our studies, however, and have not found any support for it. Several different measures of confidence and clarity of attitudes have been used. In five studies, subjects were simply asked how confident they were in their attitudes. In three, we included measures of the latitude of acceptance and latitude of rejection of people's attitudes, where they endorsed from a list of attitude statements all that they found acceptable and all they found objectionable as descriptions of their position (Sherif & Hovland, 1961). Finally, we measured how long it took subjects to answer the questions about their attitudes in two studies, with the assumption that low confidence would be reflected by longer response times.

As seen in Table VIII, no evidence has been found that thinking about reasons reduces confidence in attitudes. The only significant difference on any of the confidence measures was such that subjects who thought about reasons expressed more confidence in their attitudes than did subjects in the control condition (Wilson, 1985a; p = .01). It seems, then, that introspecting about reasons does not reduce attitude-behavior consistency by lowering confidence in one's attitudes.

C. DOE THINKING ABOUT REASONS INCREASE THE COGNITIVE DIFFERENTIATION OF ATTITUDES?

Schlegel and DiTecco (1982) have recently distinguished between attitudes that have a differentiated versus an undifferentiated cognitive component. People who are knowledgeable about or experienced with an attitude object, they argued, are likely to have complex, multidimensional beliefs that are not easily represented on unidimensional, affective scales. That is, experienced people are apt to view the attitude object as consisting of many different dimensions, some good and some bad, making it difficult to express their attitudes on a single affective scale. Their responses on such scales, Schlegel and DiTecco reasoned, will thus be poor predictors of their behavior. People with little knowledge about or experience

Schlegel and DiTecco's (1982) work provides a possible explanation of the disruptive effects of thinking about reasons. Thinking about reasons may make cognitions about the attitude object more differentiated, by forcing people to consider different aspects of the stimulus. Thus, people in the reasons-analysis conditions may end up with more differentiated attitudes than control subjects, making their responses on single-evaluative-attitude measures less predictive of

| TABLE VIII |
| THE EFFECT OF THINKING ABOUT REASONS ON CONFIDENCE IN ONE'S ATTITUDES |

<table>
<thead>
<tr>
<th>Study</th>
<th>Control condition</th>
<th>Reasons condition</th>
<th>Significance of difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Confidence questions*</td>
<td>Wilson et al. (1984)</td>
<td>5.38</td>
<td>5.48</td>
</tr>
<tr>
<td></td>
<td>Wilson (1985a)</td>
<td>7.73</td>
<td>8.20</td>
</tr>
<tr>
<td></td>
<td>Wilson et al. (1988b)</td>
<td>7.60</td>
<td>7.32</td>
</tr>
<tr>
<td></td>
<td>Wilson et al. (1988b)</td>
<td>7.56</td>
<td>7.00</td>
</tr>
<tr>
<td></td>
<td>Wilson &amp; Dunn (1986b)</td>
<td>7.00</td>
<td>7.61</td>
</tr>
<tr>
<td>Latitude of acceptance*</td>
<td>Wilson (1985a)</td>
<td>2.16</td>
<td>2.41</td>
</tr>
<tr>
<td></td>
<td>Wilson et al. (1988b)</td>
<td>2.05</td>
<td>2.63</td>
</tr>
<tr>
<td></td>
<td>Wilson et al. (1988b)</td>
<td>2.61</td>
<td>2.83</td>
</tr>
<tr>
<td>Latitude of rejection*</td>
<td>Wilson (1985a)</td>
<td>5.87</td>
<td>6.00</td>
</tr>
<tr>
<td></td>
<td>Wilson et al. (1988b)</td>
<td>6.15</td>
<td>6.05</td>
</tr>
<tr>
<td></td>
<td>Wilson et al. (1988b)</td>
<td>5.11</td>
<td>5.50</td>
</tr>
<tr>
<td>Response time to attitude questions*</td>
<td>Wilson et al. (1984)</td>
<td>4.81</td>
<td>4.78</td>
</tr>
<tr>
<td></td>
<td>Wilson &amp; Dunn (1986b)</td>
<td>8.67</td>
<td>7.92</td>
</tr>
</tbody>
</table>

*Higher numbers reflect confidence. All ratings were made on 9-point scales except for those in the Wilson et al. (1984) study, which were made on 7-point scales.

*Smaller numbers reflect greater confidence.

*Large numbers reflect greater confidence.

*p < .05.
their behavior. If so, the low attitude–behavior correlations we have found would essentially be a measurement problem, correctable if multidimensional attitude measures rather than single affective scales were used.

We suggest, however, that the type of short-term reasons analyses that people perform in our studies do not produce multidimensional attitudes. Thinking about reasons for 10 minutes is not sufficient to turn an unknowledgeable person into a knowledgeable person who has a well-integrated, differentiated belief structure. To see if thinking about reasons increases attitude differentiation, we replicated Schlegel and DiTecco’s (1982) survey, with the addition of a reasons-analysis manipulation (Wilson & Dunn, 1986b). Over 300 students were given the same survey about attitudes toward marijuana, and they were divided into knowledgeable and unknowledgeable groups, based on their reported use of marijuana. Before filling out the survey, half of the respondents were asked to describe why they liked or did not like to smoke marijuana (reasons-analysis condition), while the other half described why they chose the college they did (control condition).

Following Schlegel and DiTecco (1982) a hierarchical factor analysis was performed on subjects’ beliefs about marijuana in each of the four Knowledge × Reasons-Analysis cells. (A principal axes factor solution with oblique rotation was used to derive first-order factors. Higher-order factors were then derived by factor analyzing in the same way the correlations among the rotated first-order factors.) This analysis yielded four measures of attitude structure in each cell: (1) attitude differentiation (the number of lower-order factors resulting from the hierarchical factor analysis), (2) attitude organization (the total communality of the factor solution divided by the total normalized variance), (3) attitude centrality (the variance explained by the most general factor divided by the total communality of the solution), and (4) attitude complexity (the number of factors at each level of the solution weighted by the variance at each level, multiplied by the number of levels, divided by the degree of organization). These indices of cognitive structure are virtually identical to Zajonc’s (1965), except that Schlegel and DiTecco (1982) prefer the label “organization” to Zajonc’s label of “unity.”

As seen in Table IX, Schlegel and DiTecco’s (1982) results were replicated: People experienced with marijuana had attitudes that were more differentiated

<table>
<thead>
<tr>
<th>Condition</th>
<th>Differentiation</th>
<th>Organization</th>
<th>Centrality</th>
<th>Complexity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unknowledgeable Control</td>
<td>3</td>
<td>.624</td>
<td>.333</td>
<td>7.598</td>
</tr>
<tr>
<td>Reasons analysis</td>
<td>3</td>
<td>.621</td>
<td>.333</td>
<td>7.329</td>
</tr>
<tr>
<td>Knowledgeable Control</td>
<td>4</td>
<td>.581</td>
<td>.250</td>
<td>8.957</td>
</tr>
<tr>
<td>Reasons analysis</td>
<td>4</td>
<td>.579</td>
<td>.250</td>
<td>9.176</td>
</tr>
</tbody>
</table>

and complex and less central than people inexperienced with marijuana. (The organization index revealed few differences between knowledgeable and unknowledgeable subjects in either our study or Schlegel and DiTecco’s.) The results show no evidence, however, that thinking about reasons increases the dimensionality of attitudes: The measures of differentiation, organization, centrality, and complexity were virtually identical in the control and the reasons-analysis conditions. Further, there was no evidence in this study or in two others (Wilson, Kraft, & Dunn, 1988a, Study 2; Wilson et al., 1984, Study 3) that the behavior of people who think about reasons can be predicted from multidimensional measures of attitudes. These results suggest that the disruptive effect of thinking about reasons on attitude–behavior consistency is not a measurement problem that can be corrected with the inclusion of multidimensional attitude scales.

D. DEMAND CHARACTERISTICS AND SELF-PRESENTATIONAL CONCERNS

Perhaps asking people to think about reasons heightens self-presentational concerns, compelling people to report an attitude they thought was socially acceptable, but which they knew was not their true attitude. Alternatively, people may have felt an implicit demand to report a new attitude after thinking about reasons. In either case, attitude–behavior correlations would suffer, because behavior (which subjects did not know was being assessed) would reflect people’s true attitudes, whereas their reports would reflect attitudes that they wanted or thought they were expected to hold.

Tetlock and Manstead (1985) have cogently pointed to the difficulties of distinguishing between impression management versus intrapsychic explanations. It is relatively clear, however, that neither concerns about one’s public image nor demand characteristics were responsible for the changes in reported attitudes in people who thought about reasons. In all of our studies (with the exception of Wilson and Kraft’s [1988] phone survey) subjects’ attitude reports were made privately and anonymously. In one study, subjects never marked their responses on paper; they pressed buttons that supposedly sent their responses directly to a computer that immediately aggregated all subjects’ responses (Wilson et al., 1984, Study 1). Thus, there is every reason to believe that subjects were candid in their responses.

Nor is it likely that demand characteristics can account for our results. In one study, subjects thought their reasons-analysis, attitude, and behavioral responses were part of three separate studies (Wilson & Dunn, 1986a, Study 1), thus it is unlikely that they viewed the reasons-analysis manipulation as an attempt to change their attitudes. In addition, subjects were told in most of our studies that they would not hand in the paper on which they wrote down their reasons, as its only purpose was to organize their thoughts. Typically, when the experimenter
enters the room after subjects have completed the reasons questionnaire, he or she casually deposits the questionnaire in a trash can, saying that it will not be needed anymore. Further, in one study, the subjects never wrote down what their reasons were (Wilson et al., 1984, Study 2).

Perhaps the best evidence against a demand characteristic or self-presentational interpretation comes from a measure used in the Wilson, Lisle, and Schooler (1988c) study. In this study, the attitude object was a set of five posters commonly purchased by college students. At the end of the study, subjects were allowed to choose one of the posters to take home, and they thought the experimenter would not know which one they took. If subjects who thought about reasons reported new attitudes purely for self-presentational reasons or due to demand characteristics, one would not expect this change to show up on the measure of which poster they chose to take home. (Because they thought they would not know which poster they chose, they presumably chose the one they genuinely liked the best.) Not only did subjects in the reasons condition change their minds about which posters they liked the most, however; they also tended to choose different posters to take home. Thus, demand characteristics or self-presentational concerns appear not to have played a major role in the attitude change that occurred in this study.

VIII. Conditions under Which People Think about Reasons

An important issue in research on analyzing reasons concerns the conditions under which a reasons analysis is performed. In our studies, people were instructed to think about reasons, and the effects of this reasons analysis were then examined. This manipulation has ecological validity, in that it is fairly common in everyday life to be asked to explain one’s feelings (“Why do you like her? Why didn’t you like the movie?”). Nonetheless it is important to examine other conditions triggering a reasons analysis. We do not go through our daily lives constantly examining our motives and reasons; indeed, Langer (1978) argued that much of the time, we follow well-learned scripts in a relatively mindless fashion. When are people likely to think about the reasons for their feelings?

Though there has been little research on this specific question, the issue of when people attempt to explain other people’s behavior has received a considerable amount of attention. Causal attributions about other people are apt to be made when (1) people are explicitly asked to make attributions, as in our reasons-analysis studies; (2) an unexpected event occurs, such as when a friend behaves in an uncharacteristic manner (Clary & Tesser, 1983; Hastie, 1984; Lau & Russell, 1980; Pyszczynski & Greenberg, 1981; Wong & Weiner, 1981); (3) a stimulus person has hedonic relevance to the observer (Berscheid, Graziano, Monson, & Dermer, 1976; Harvey, Yarkin, Lightner, & Town, 1980; Monson, Keel, Stephens, & Genuing, 1982), and (4) people experience a lack of control (Pittman & Pittman, 1980).

The factor that has received the most attention, and seems to be most predictive of whether causal attributions about other people are made, is the unexpectedness of someone else’s behavior. This same factor might also determine when an analysis of reasons for our own behavior and feelings occurs. People have many expectations about how they will feel toward an attitude object; for example, most people fully expect to enjoy the next chocolate bar they eat, and they expect their next trip to the dentist to be unpleasant. If their reactions to future chocolate bars and trips to the dentist were consistent with their expectations, there is little reason to think about why. If, however, their reactions violated their expectations (a scintillating dentist appointment?), they are likely to try to explain these reactions.

If having unexpected feelings triggers a reasons analysis, and the other conditions for the disruptive effects of thinking about reasons are met (such as the attitude being affective), then the following sequence of events should occur: (1) People notice that their reaction to a stimulus is not what they expected it to be, (2) they attempt to explain their unexpected reaction, (3) they recast their attitude in cognitive terms, resulting in low attitude–behavior correlations.

We have conducted three experiments to test this hypothesized sequence of events, as part of a research program on the effects of expectations about one’s own affective reactions (Wilson, Lisle, Kraft, & Wetzel, in press). We describe portions of this research that bear directly on the hypothesis that having unexpected affective reactions triggers a reasons analysis. Rather than instructing people to think about reasons in these studies, we manipulated people’s expectations about how they would feel toward an attitude object, then we violated these expectations in some conditions, with the assumption that a reasons-analysis would be triggered. Two control conditions were included in each study: (1) One group was given expectations about how they would feel, just as in the unexpected reaction condition, but these expectations were confirmed rather than disconfirmed; (2) another control group was given no expectations about how they would feel. The main dependent measure was, as in most of our previous studies, the correlation between subsequent attitudes and behavior. To check our assumption that having unexpected feelings triggers a reasons analysis, subjects in all conditions completed a thought-listing questionnaire at the end of the experiment (Brock, 1967; Cacioppo & Petty, 1981; Greenwald, 1968), on which they listed all the thoughts they could recall having when they were first exposed to the attitude object.

In Study 1, for example, subjects were asked to taste a beverage, and they were told that a very high percentage of people had liked the beverage. Some subjects received a pleasant-tasting lemonade—thus there was no discrepancy between their affective reaction and their expected reaction. Other subjects
received a fairly unpleasant concoction of lemonade, water, and salt—thus there was a discrepancy between their affective reaction and their expected reaction. Finally, a no-expectation control condition was included in which subjects were not told about other subjects’ reactions, with half of the control subjects receiving the pleasant drink and half the unpleasant drink. The attitude measure in this study was subjects’ rating of how much they enjoyed the beverage, whereas the behavioral measure was the amount of the beverage subjects drank during a free-time period when they thought the study was over.

Two other studies have also been conducted to assess the effects of violating affective expectations, using (1) different attitude objects (single-panel cartoons and reproductions of modern art) and (2) a different way of manipulating people’s expectations. In Study 2, subjects viewed and rated 20 cartoons, and they received feedback indicating that the ones they disliked were all from the same magazine. They then saw four new cartoons that were supposedly from the same magazine, but were (according to pretest subjects) fairly funny. Thus, they expected to dislike the new cartoons, but these expectations were disconfirmed. Subjects in the expected-reaction condition received feedback indicating that the ones they had liked from the first set were from the target magazine—thus they expected to like the four new cartoons. Subjects in the no-expectation condition did not see any cartoons from the target magazine before viewing the four new ones. The attitude measure was subjects’ ratings of how funny the new cartoons were, while the behavioral measure was the length of time they spent viewing the new cartoons.

Study 3 had the same design as Study 2, except that subjects viewed reproductions of paintings rather than cartoons. They saw new paintings at the end of the study that were ostensibly all of the same (fictional) style, labeled the Schematic Style. As in Study 2, subjects initially rated 20 paintings, and some subjects received feedback indicating that the ones they liked were all of the Schematic Style. Others received feedback that the ones they disliked were of the Schematic Style, while the remainder were not told that any of the first 20 were Schematic paintings. The new Schematic paintings all subjects subsequently saw had been given low ratings by pretest subjects, thus some subjects had unexpected reactions to them, others had expected reactions, and some had no expectations about how they would react. There was no behavioral measure in this study, thus attitude–behavior correlations could not be assessed. The thought-listing task was included, to see if having unexpected reactions triggered a reasons analysis.

Our main hypothesis was that subjects in the unexpected-reaction conditions would think about why their reactions were not what they expected them to be, and that this reasons analysis would result in low correlations between their attitudes and behavior. Subjects in the expected-reaction conditions should respond similarly to those with no expectations. Since their feelings toward the attitude objects were consistent with their expectations, there was no reason to think about why—thus, no reasons analysis was expected to occur, and attitude–behavior correlations were expected to be high.

As seen in Table X the predicted pattern of attitude–behavior correlations occurred: They were high in the no-expected and expected-reaction conditions, and low in the unexpected-reaction conditions. Averaging across studies, the correlation in the unexpected reaction condition was significantly lower than both the correlation in the no-expectation condition and the correlation in the expected-reaction condition ($p'\leq.05$).

Subjects’ responses on the thought-listing task were coded in each study to test the hypothesis that the lowered correlations in the unexpected-reaction conditions were caused by a reasons analysis. As it happened, the coders achieved acceptable reliability only on a rather broad categorization of the reported thoughts, namely a division of the thoughts into those that concerned the attitude object (e.g., the beverage in Study 1) and those that did not. The number of thoughts in each category, as well as the total number of thoughts reported, are displayed in Table X. As predicted, subjects in the unexpected-reaction conditions reported the most thoughts, while subjects in the no-expectation and expected-reaction conditions reported fewer similar numbers of thoughts. The total number of thoughts reported by subjects in the unexpected-reaction condition was significantly higher than the total number of thoughts reported in each of the other two conditions ($p'\leq.05$).

To summarize the results of these studies, the last part of our prediction—that having unexpected reactions triggers a reasons analysis, which leads to low attitude–behavior consistency—has been supported. We also obtained evidence that having unexpected reactions triggers more thoughts than having expected

<table>
<thead>
<tr>
<th>Condition</th>
<th>No expectation</th>
<th>Expected reaction</th>
<th>Unexpected reaction</th>
</tr>
</thead>
<tbody>
<tr>
<td>Attitude–behavior correlations</td>
<td>69.7</td>
<td>70.1</td>
<td>.112</td>
</tr>
<tr>
<td>Thoughts about the attitude object (.86)</td>
<td>3.221</td>
<td>3.241</td>
<td>4.051</td>
</tr>
<tr>
<td>Thoughts not about the attitude object (.72)</td>
<td>1.571</td>
<td>1.557</td>
<td>2.087</td>
</tr>
<tr>
<td>Total thoughts (.89)</td>
<td>4.797</td>
<td>4.781</td>
<td>6.124</td>
</tr>
</tbody>
</table>

*The correlations were averaged over two studies, the reported thoughts over three studies. Correlations or means in the same row that have different superscripts, 1 and 2, differ from each other at the .05 level of significance. The numbers in parentheses after each category are the reliabilities of the coders’ ratings, averaged across studies.
reactions or having no expectations. The evidence for the hypothesis that this increase in thinking involves more of a reasons analysis, however, could not be tested because our coders were unable to achieve sufficient reliability when attempting to categorize the thoughts as reasons. The results of our initial work in this area are promising, however, and suggest that further tests of the hypothesis be performed, with more reliable measures of people’s tendency to think about reasons.5

It is important to reiterate that these studies are also useful in ruling out a demand characteristic or self-presentation interpretation of our earlier experiments, in which subjects were asked to think about their reasons for liking or disliking an attitude object. Despite our efforts to convince subjects that their reasons analyses were private and that no one would ever see their reasons, it is possible that they were still concerned with putting their best reasons forward. Further, simply asking subjects to think about reasons might imply to them that they should rethink their position and possibly change their minds. In the set of studies just reviewed, however, subjects were never asked to think about reasons. Instead, the situation was set up so that a reasons analysis should be triggered naturally (i.e., by manipulating the unexpectedness of subjects’ affective reactions). As predicted, this both increased the number of thoughts subjects had and lowered their attitude–behavior consistency.

IX. How and When Will Thinking about Reasons Get Us into Trouble?

We began this article with a discussion of introspection as a uniquely human trait. In the course of our review, we have seen that one type of introspection—thinking about the reasons for one’s feelings—can change people’s minds about how they feel and lower the consistency between their attitudes and subsequent behavior. We have used the word “disruptive” to describe these effects, in part because we wish to avoid a general condemnation of introspection. We do believe, however, that it is not always to people’s advantage to turn an affective response into a cognitive one, as seems to occur as a result of thinking about reasons. In our next, concluding, section, we discuss further when it is advantageous and when it is not to think about the reasons for one’s feelings. In this section, we describe two empirical demonstrations that the attitude change that results from thinking about reasons can be disadvantageous to the person doing the introspecting.

A. THE FUNCTION OF ATTITUDE OBJECTS

The primary function of many attitude objects is to give pleasure. When deciding what flavor of ice cream to buy, whether to paint the living room yellow or blue, or whom to date on Saturday night, the overriding goal is to make the choice that is most affectively pleasing. It is precisely these types of attitude objects, we suggest, about which one should avoid thinking about reasons. If we spend too much time thinking about why we prefer vanilla to rocky road, why yellow seems preferable to the blue, or why we like Bob more than Joe, we may recast our feelings in cognitive terms, and thereby lose sight of what our preferences are.

Wilson, Lisle, and Schooler (1988a) performed two experiments to test this hypothesis. Two conditions needed to be satisfied in these studies: (1) Stimuli needed to be used whose function was primarily affective—that is, people’s main criterion for choosing among them is the extent to which they are affectively pleasing; (2) a criterion needed to be found for assessing the validity of people’s preferences or choices. To demonstrate that thinking about reasons can be undesirable, a standard of judgment needs to be found by which people’s preferences can be evaluated.

In Wilson, Lisle, and Schooler’s (1988a) first study, the attitude object was a food item—various kinds of strawberry jams. This seems to satisfy the criterion that it be a stimulus whose function is largely affective. The standard of judgment we used was evaluations of the jams made by taste experts at Consumer Reports magazine. These taste experts rated 44 jams according to 16 sensory attributes (e.g., aroma, sweetness, bitterness), and the jams were ranked ordered by Consumer Reports on the basis of these ratings.

For our study, we chose five of the jams that were wide apart in their rankings by the Consumer Reports experts (ranging from the first to the last). Subjects tasted a sample of each of the jams and rated how much they liked it, ostensibly as part of a consumer psychology experiment. We assumed that in the absence of a reasons analysis, subjects would rate the jams according to their affective reaction toward each one, and that these affective reactions would correspond well to the ratings of the Consumer Reports taste experts. This is what happened; the average within-subject rank-order correlation between control subjects’ liking ratings and the ranking of the taste experts was .55, which was significantly greater than zero.

Half of the subjects received our standard reasons-analysis manipulation, in
change people's minds about how they felt, presumably by turning an affective response into a more cognitive one. This prediction was confirmed: Subjects in the reasons condition rated the art posters lower and the popular posters higher than did control subjects, and they were significantly more likely to choose one of the popular posters to take home.

These results still do not demonstrate that the change in attitudes that occurred were in any way disadvantageous for subjects in the reasons condition. As in the jam study, there may not be any undesirable consequences of changing one's mind about which attitude object is the best. A follow-up measure, however, suggested that there are such negative consequences to thinking about reasons. Subjects were telephoned approximately 3 weeks after they were in the study and asked a series of questions about their satisfaction with their choice of poster (e.g., whether they still had the poster and whether they had hung it up). Subjects in the reasons condition indicated that they were significantly less pleased with their choice than did subjects in the control condition, suggesting that an initial affective reaction to the poster (which, as indicated by control subjects, was most positive toward the art posters) had had a chance to reassert itself. As discussed earlier, the results of the poster study were moderated by people's knowledge about art. As predicted, they were stronger: for subjects who were unknowledgeable.

This result is important, because it suggests that people ought not to be too introspective about many personal and consumer decisions, particularly if they are not very knowledgeable in that domain. Trying to think about the reasons for one's preferences can change people's minds about how they feel, leading to a choice that they later regret. In other words, some things are best left unanalyzed. This view was well-stated in an article by Anna Quindlen, in which she discusses her attitude toward rock and roll:

Some people overanalyze rock-and-roll, just as they overanalyze everything else. They say things like "Bruce Springsteen is the poet laureate of the American dream gone sour," when all I need to know about Bruce Springsteen is that the saxophone bridge on "Jungleland" makes the back of my neck feel exactly the same way I felt the first time a boy kissed me. . . . Rock-and-roll is a lot like sex: If you talk seriously about it, it takes a lot of the feeling away—and feeling is the point. (Quindlen, 1987, p. C12)

B. STIMULI WITH AFFECTIVE AND PRACTICAL FUNCTIONS

It is important to note that the two Wilson, Lisle, and Schoolder (1988c) studies used stimuli with a primarily affective function. With stimuli such as jams and posters, there is little else to consider other than the extent to which they give pleasure. That is, people need only heed the affective component of their attitudes, and they probably should avoid forms of introspection that emphasize cognitions
about the attitude object. For many other kinds of stimuli, however, both affective and practical considerations are important, and we may be better advised to heed both the affective and cognitive components of our attitudes.

For example, consider the decision of which car to buy. Clearly, there is an affective component to this choice. People’s attachment to their automobile has been described as “truly passionate,” stemming from “a deep-seated desire to include the car in the family and to emphasize our strong, symbolic attachment with these objects of affection” (Marsh & Collett, 1987, p. 18). Unlike jams or posters, however, there are many attributes of cars other than their affective appeal that probably ought to be considered, such as their price, safety, frequency of repair, and fuel consumption. In fact, many of the major decisions we make in life—who to marry, what job to accept, what house to buy—involve a weighting of both the affective and cognitive components of our attitudes toward the alternatives.

An interesting area of future research would be to examine the effects of thinking about reasons on decisions involving stimuli such as these. Because the cognitive component of attitudes is important in these decisions, introspection (such as analyzing reasons) that emphasizes this component may be useful. If thinking about reasons obscures people’s affect too much, however, they may well end up making a decision they later regret. It appears that a balance between thinking about reasons and “listening to one’s heart” is the best strategy for these types of decisions.

X. Summary and Conclusions

A. UNANSWERED QUESTIONS

Throughout this article, we have pointed out areas in which further evidence is needed to nail down a particular hypothesis or conclusion, and there is no need to repeat these comments here. One important unanswered question, however, bears mentioning. It is clear that asking people to think about reasons will often produce attitude change, particularly for affectively based attitudes. The direction of this change, however, has been difficult to predict. In some of our studies, people who think about reasons end up with an attitude that is significantly more negative or positive, on the average, than the attitudes of control subjects. In the Wilson, Lisle, and Schooler (1988c) poster study, for example, subjects who thought about reasons reported greater liking for the popular posters and less liking for the art posters than did control subjects. In the Wilson and Kraft (1988) study that manipulated the basis of attitudes, subjects who were in the mere-exposure condition became more negative toward the clinical psychology students than control subjects. In other studies, thinking about reasons causes changes in both a positive and a negative direction, averaging out to no difference when compared to control subjects. This result was found in Wilson, Kraft, and Dunn’s (1988a) study on attitudes toward the 1988 Presidential candidates, and in Wilson and Kraft’s (1988) survey of attitudes toward social issues. (Presumably, this bidirectional change also occurred in Wilson et al.’s [1984] couples study.

We suggest that the reason the direction of attitude change is difficult to predict is closely related to our hypothesis about the generation of a biased sample of reasons. We have argued that people’s reasons are often biased in the direction of such factors as what seems most plausible as a cause, what is easiest to verbalize, and what is most available in memory. Because it is difficult to know what is available in memory for any given subject at any given time, it is not surprising that the direction of the resulting attitude change has been difficult to predict. In the Wilson et al. (1984) couples study, for example, we had no way of knowing whether positive or negative reasons were most available to any given couple.

It would be possible, of course, to predict the direction of attitude change by manipulating the types of reasons that are available to subjects. This is precisely what was done by Salancik (1974) and by Seligman et al. (1980), by means of the wording of the questions subjects were asked about their reasons. As predicted, subjects became more positive toward the attitude object when intrinsic reasons were elicited and more negative when extrinsic reasons were elicited. Thus, if thinking about reasons is ever to be used as an attitude-change technique, those who use it should attempt to manipulate (or at least predict) the types of reasons subjects bring to mind, in order to know the direction of any attitude change that occurs.

B. RECOMMENDATIONS

We are now in a better position to recommend when people should introspect about the reasons for their feelings and when they should not (qualified, of course, by the proverbial call for further research). First, our data imply that it might not be good to think about reasons to the extent that we value consistency between attitudes and behavior. This is a complicated notion, however, for several reasons. First, it is clear that thinking about reasons will not always reduce attitude–behavior consistency. Some of the boundary conditions that have been established empirically are that analyzing reasons has little effect on people knowledgeable about the attitude object (presumably because they have better access to the actual causes of their attitudes), and analyzing reasons does not loosen the ability of attitudes to predict cognitively based behavior (Millar & Tesser, 1986). Second, we presented evidence that people do respond consistently with their new attitudes...
immediately after thinking about reasons, before their affective reaction to the
stimulus has a chance to reassert itself.

Even when thinking about reasons does reduce attitude–behavior consistency,
this outcome is not always to be avoided. To assess the desirability of analyzing
reasons, we must consider for whom it is desirable or undesirable—the person
doing the introspection or the social scientist attempting to predict behavior—as
well as the nature of the attitude that is being analyzed.

1. Desirability to the Social Scientist of Asking People
to Think about Reasons

There is little doubt that it can be undesirable for social scientists to ask their
respondents to think about reasons. The goal of a good deal of research in the
social sciences—including psychology, sociology, economics, political science,
and consumer research—is to predict people’s behavior from their verbally
reported attitudes. Occasionally, in these areas of research, people are asked to
explain why they feel the way they do. If “why” questions are asked before
people report their attitudes, attitude–behavior consistency will often suffer. For
example, if people are asked to think about why they prefer a particular consumer
good before reporting their preferences, their preferences may not predict their
future purchasing behavior. Unless consumers have a well-developed cognitive
structure about the product (which is unlikely when a new product is being test
marketed), asking them to justify their preferences in cognitive terms will hinder
the goal of marketing researchers—the prediction of consumer behavior.

Similarly, if members of the electorate have attitudes toward a candidate that
are primarily affectively based, asking them to think about the reasons for their
feelings could lower the correlation between their attitudes and their voting
behavior. The use of “why” questions may be justified more easily when researchers
are certain that the attitudes they are assessing are the result of a conscious scrutiny
of the issues, so that people can accurately report why they feel the way they do.

2. Desirability to the Individual Who Thinks about
Reasons

As we have seen, whether or not a reasons analysis is desirable to the person
doing the analyzing depends on the nature of the attitude under examination.
Several of our studies have shown that if the attitude is based on a good deal of
knowledge about the attitude object, then thinking about reasons will not change
people’s attitudes or lower attitude–behavior consistency. Wilson and Kraft’s
(1988) mere-exposure study suggested that this is because knowledgeable people
are more apt to have attitudes that are based on cognitive information that is
easy to verbalize, whereas people who are unknowledgeable are more apt to
have affectively based attitudes.

When an attitude is affectively based, and when the function of the attitude
object is primarily affective, the Wilson, Lisle, and Schoolder (1988c) studies
demonstrated that it is best to avoid a reasons analysis. The preferences of subjects
who thought about why they liked certain strawberry jams changed in a direction
away from an objective standard of quality, and subjects who thought about why
they liked or disliked a set of posters were more apt to regret their decision as
to which poster to take home. Even when an attitude is affectively based, however,
we must qualify our conclusion that it best not to try to explain it. First, we
noted that for many important decisions, such as what car to buy or whom to
marry, it is probably best to weigh both one’s affect toward the alternatives and
one’s beliefs about their attributes. Here, people probably ought to strike a balance
between listening to one’s heart and being very deliberative and introspective.
Exactly how to accomplish this—or even whether such a balancing is desirable
for these types of decisions—is an important avenue of future research.

Further, some affectively based attitudes are themselves undesirable (either to
the people who hold them or to the people around them); thus it may be good
to disrupt them by thinking about reasons. Our argument here is that analyzing
reasons is an attitude-change technique that may be useful in changing undesirable
attitudes. It is difficult, as discussed by Zajonc and Markus (1982), to change
an affectively based attitude by presenting people with logical, persuasive argu-
ments. If people with such attitudes were asked to explain why they feel the
way they do, however, the act of constructing an explanation might produce
some attitude change. It is unlikely that thinking about reasons is a powerful
way to change attitudes; in some of our studies, the amount of change was small,
and the extent to which this change persists over time is not clear. Nonetheless,
it may be a useful way of unfreezing an affectively based attitude by making
people reconsider it in cognitive terms.

There are at least two sorts of harmful attitudes that might be changed by
thinking about reasons—those that are undesirable from the individual’s perspec-
tive, and those that are undesirable from a societal perspective. Either type
of attitude might change in a desirable direction if people attempt to justify it.
For example, Tesser, Leone, and Clary (1978) asked people who were anxious
about speaking in public to think about why they felt the way they did when
giving a speech, and then they asked subjects to give a brief speech. Subjects
who thought about reasons reported feeling significantly less anxious during their
speech than did subjects in a control, no-introspection condition. Thus, a feeling
that was undesirable from the subjects’ perspective—speech anxiety—was
changed by asking people to justify it (see also Leone & Baldwin, 1983; Leone,
Minor, & Baltimore, 1983). When people think about their reasons for their
anxieties, they presumably find it difficult to think of logical, rational reasons,
and this inability to justify their negative affect succeeds in changing it. It is not
clear how long-lasting such change is, thus we are not suggesting that thinking
about reasons be adopted as a powerful therapeutic technique. Nonetheless, the
Tesser et al. (1978) study is a good example of a case in which analyzing reasons
can have beneficial rather than undesirable consequences.

Analyzing reasons may also be useful as a way of challenging attitudes that
are harmful from a societal perspective. Generalizing from the Tesser et al. (1978)
study, people asked to justify their racist attitudes might find it difficult to do
so, and thus they might moderate their beliefs. Again, we are not suggesting
that this will necessarily produce long-lasting reductions in prejudicial attitudes;
indeed, the available evidence suggests that changes in attitudes will be more
likely to occur than changes in behavior. With undesirable feelings such as racial
prejudice, however, it is desirable to unfreeze the attitude, even if behavioral
change does not immediately follow.

It is best not to think about reasons when the function of the attitude object
is to give us pleasure, when our attitude is affectively based, and when we have
desire to change our feelings. If we couple is quite happy with their feelings
toward each other, attempts to justify these feelings may not be productive. Many
affectively based attitudes cause no discomfort (unlike speech anxiety) and cause
no harm to others (unlike racial prejudice), including feelings about our spouses,
friends, and hobbies—not to mention strawberry jams and art posters. When
asked to report these kinds of feelings, people would do well to follow the advice
one of us recently received in a fortune cookie: “Answer just what your heart
prompts you.”

Acknowledgments

The writing of this chapter and the research it reports were supported by grants from the National
Science Foundation (BNS-8316189) and the National Institute of Mental Health (MH41841). We
would like to thank several people who read a previous draft of this chapter and who provided
extremely valuable feedback, including Robert Abelson, Gene Borgida, Bella DePaulo, Arie Krug-
lanski, Richard Petty, Jonathan Schooler, Jeanne Smith, Mark Snyder, Claude Steele, Abe Tesser,
and Chris Wetzel.

References

components in political person perception. *Journal of Personality and Social Psychology, 42*,
619–630.

Berscheid, E., Graziano, W., Monson, T., & Dermer, M. (1976). Outcome dependency: Attention,

Brock, T. C. (1967). Communication discrepancy and intent to persuade as determinants of coun-

assessment: The thought-listing technique. In T. Merluzzi, C. Glass, & M. Genest (Eds.),

York: Wiley.


Carper, J., & Doob, L. W. (1953). Intervening responses between questions and answers in attitude

to human behavior. *New York: Springer-Verlag*.

Clary, E. G., & Tesser, A. (1983). Reactions to unexpected events: The naive scientist and interpretive

(Ed.), *The quantitative analysis of social problems* (pp. 168–189). Reading, MA: Addison-
Wesley.

Press.


215–251.

MA: MIT Press.

The handbook of motivation and cognition: Foundations of social behavior (pp. 204–243). New
York: Guilford Press.

behavior consistency, and the strength of the object–evaluation association. *Journal of Exper-


In A. Greenwald, T. Brock, & T. Ostrom (Eds.), *Psychological foundations of attitudes* (pp.


Psychology, 46*, 44–56.


Hovland, C. I. (1959). Reconciling conflicting results derived from experimental and survey studies


CONTENTS

Contributors.............................................................................. ix

On the Construction of the Anger Experience: Aversive Events and Negative Priming in the Formation of Feelings
Leonard Berkowitz and Karen Heirzer

I. Introduction ......................................................................... 1
II. Approaches to Emotion and Anger ....................................... 2
III. A Cognitive–Neoassociationistic Analysis of Anger ............... 7
IV. The Present Research ....................................................... 13
V. General Discussion .......................................................... 31
References ............................................................................. 35

Social Psychophysiology: A New Look
John T. Cacioppo, Richard E. Petty, and Louis G. Tassinary

I. Introduction ......................................................................... 39
II. Background ........................................................................ 40
III. An Alternative Conceptualization ................................. 43
IV. Illustrating the New Look: Part I .................................... 43
V. Illustrating the New Look: Part II .................................... 65
VI. Sociological and Philosophical Obstacles ..................... 73
VII. The New Look: Part III ................................................. 77
VIII. Conclusion ....................................................................... 81
References ............................................................................. 83

A Self-Discrepancy Theory: What Patterns of Self-Beliefs Cause People to Suffer?
E. Tory Higgins

I. Introduction ......................................................................... 93
II. Self-Discrepancy Theory .................................................. 94