Effects of Introspecting About Reasons: Inferring Attitudes From Accessible Thoughts

Timothy D. Wilson, Sara D. Hodges, and Suzanne J. LaFleur
University of Virginia

Previous research has shown that analyzing reasons can change people's attitudes, but the exact mechanisms of this effect have not been entirely clear. It was hypothesized that introspecting about reasons focuses people's attention on thoughts that are accessible in memory and increases the extent to which people view their accessible thoughts as applicable to their current attitudes. In Study 1, college students formed initial impressions of a target person, and then positive or negative thoughts about the target person were made memorable. After a delay, half of the participants analyzed reasons for their attitude and half recalled the target person's behaviors. As predicted, people who analyzed reasons reported attitudes toward the target person that were based more on what they could recall about her. Study 2 showed that this effect occurs regardless of whether people initially form an on-line impression. Implications for the effects of analyzing reasons and for attitude formation are discussed.

Think, for a moment, about an acquaintance you have recently met. Why do you feel the way you do about this person? A considerable amount of research indicates that answering this question can change your attitude toward your acquaintance. Thinking about reasons has been found to change people's attitudes toward a variety of attitude objects, including other people, food items, puzzles, and works of art (e.g., Johnson, Mac-Arthur, & Wright, 1991; Millar & Tesser, 1986, 1989; Wilson, Bybee, Dunn, Hyman, & Rotondo, 1984; Wilson & Dunn, 1986; Wilson & Kraft, 1993; Wilson, Kraft, & Dunn, 1989; Wilson et al., 1993; Wilson & Schooler, 1991). Wilson and Kraft (1993), for example, asked college students involved in steady dating relationships to think about why they felt the way they did about their relationships. In comparison with members of a control group who did not analyze reasons, these students were significantly more likely to change their minds about how they felt about the relationship.

Such attitude change occurs, we suggest, because of the following sequence of events: When thinking about reasons, people focus on attributes of the attitude object that are accessible in memory, plausible as causes of their feelings, and easy to verbalize. Because people do not have perfect access to the reasons for their attitudes (Nisbett & Wilson, 1977; Wilson & Stone, 1985), however, the reasons that are plausible, accessible, and easy to verbalize are often unrepresentative of the actual causes of their attitudes and, thus, can imply a somewhat different attitude than they held before they thought about reasons. People assume that the reasons they generate reflect their attitudes, much like a self-perception effect (Bem, 1972). That is, they base their attitude on the reasons that come to mind, leading to attitude change.

We have obtained a fair amount of circumstantial evidence for this proposed sequence of events. For example, the reasons people generate often imply a somewhat different attitude than they expressed earlier, and these reasons are highly correlated with the attitude they report after analyzing reasons (for reviews, see Wilson, 1990; Wilson, Dunn, Kraft, & Lisle, 1989; Wilson & Hodges, 1992). These findings suggest that people bring to mind reasons that are not entirely consistent with their initial attitude and base their subsequent attitude on these reasons. However, these findings are correlational in nature. To date, there is no direct evidence for the crux of our argument, namely that analyzing reasons focuses people's attention on thoughts about the attitude object that are most accessible in memory and that people base their subsequent attitude on these accessible thoughts. The purpose of the present studies was to provide such evidence. We asked people to form an impression of a target person and then manipulated whether positive or negative thoughts about this person were accessible in memory. We predicted that when people analyzed the reasons why they felt the way they did about this person, they would focus on the thoughts about the person that were most accessible in memory and would infer that these accessible thoughts reflected their attitude.

Our hypotheses about the effects of analyzing reasons can be understood in terms of recent research on accessibility and the "judged usability" of accessible information (Higgins, in press). As noted by several researchers, if information is to be used in social judgments, it is not sufficient that it be accessible in memory. People must also view the information as applicable to the judgment (Bem, 1972; Higgins, 1989, in press; Martin & Achee, 1992; Schwarz & Bless, 1992; Strack, 1992). For exam-

ple, Bem (1972) argued that people will infer their attitude from observations of their behavior only if the behavior is viewed as a "tact" (i.e., only if their behavior is viewed as applicable to the question "How do I feel?"). Similarly, Higgins (in press), Martin and Achee (1992), Schwarz and Bless (1992), and Strack (1992) have argued that accessibility is but one step in the process of social judgment; people also assess the appropriateness or validity of accessible information to the judgment at hand.

Our view of the effects of analyzing reasons is compatible with these arguments concerning accessibility and judged applicability. As mentioned, we have hypothesized that two steps often occur when people analyze reasons: A biased set of thoughts about the attitude object becomes accessible, and people infer that these reasons reflect their current attitude. The first step of this process involves accessibility; as a result of the limits of introspection and memory, the reasons that are accessible are often unrepresentative of people's initial attitude. The second step can be viewed as the process of judged usability noted by Bem, Higgins, and others; people often assume that their accessible reasons are applicable in the sense that they infer that these accessible reasons reflect their current attitude.

Under some circumstances, people might recognize that their reasons are not representative of their attitude and thereby avoid the inference that their reasons reflect how they feel. A study conducted by Hodges and Wilson (1994) suggested one such circumstance: the accessibility of people's initial attitude. Hodges and Wilson compared the effects of analyzing reasons among people whose initial attitudes were accessible versus nonaccessible. Interestingly, both groups were equally likely to bring to mind reasons that were somewhat inconsistent with their initial attitude: Having an accessible initial attitude did not prevent people from thinking about reasons that conflicted, at least to some degree, with their initial attitude. People with accessible attitudes, however, were significantly less likely to infer that their current attitude matched their reasons. Having an accessible attitude seemed to short-circuit the self-perception process, just as Bem (1972) proposed. In other words, people with accessible attitudes were less likely to view their accessible reasons as applicable to their current attitude than were people with inaccessible attitudes.

What about when an initial attitude is not highly accessible? Are there conditions under which people recognize that the reasons that come to mind are biased and thus not very good indicators of how they currently feel? In other domains, people have been shown to avoid biased judgments, at least to some extent, when it is obvious that the information they have obtained is tainted or invalid. Examples include instances in which jurors are made aware of the potentially biasing effect of certain kinds of testimony and when people are made aware of the relevance and applicability of statistical information (e.g., Kruglanski, Friedland, & Farkash, 1984; for a review, see Wilson & Brekke, 1994). Similarly, when people in priming studies are made aware that their accessible thoughts were caused by an arbitrary priming manipulation, they are less likely to assimilate their judgments to these thoughts (for reviews, see Higgins, in press; Martin & Achee, 1992; Schwarz & Bless, 1992; Strack, 1992). Perhaps there is a parallel in the domain of attitudes: If it is obvious that one's thoughts about the attitude object are biased

or skewed, one may be able to avoid the inference that these thoughts reflect one's attitudes.

A study conducted by Greenwald (1968) is consistent with this hypothesis. He asked participants to read six arguments for and six arguments against the question of whether the United States should give aid to foreign countries. After reading these arguments, half of the participants were asked to rehearse only the pro arguments, whereas the others were asked to rehearse only the con arguments. One week later, people remembered more of the arguments they had been asked to rehearse than those they had not. Interestingly, however, there was no difference in people's attitude toward foreign aid between the different conditions. People seemed to have formed their attitude on the basis of their initial reading of the arguments and did not base their attitude on the arguments that were accessible in memory, perhaps because they recognized that what was most memorable was biased as a result of the experimenter's instructions to rehearse only the pro or con arguments.

Are there similar circumstances under which people asked to analyze reasons become aware of the fact that their accessible reasons are biased (i.e., not representative of their attitudes) and thereby do not infer that these reasons reflect their current attitudes? In contrast to Greenwald's (1968) study, we suggest that it is very difficult for people to recognize the invalidity of their self-generated reasons. First, when people think about why they feel the way they do, the source of bias on their reasons is not nearly so obvious as in the Greenwald (1968) study, in which the experimenter asked people to recall a subset of arguments. In our studies, there is no external source of bias on people's reasons; we simply ask participants why they feel the way they do. People's reasons often are constrained by the limits of human introspection, and yet people do not seem to recognize this fact. That is, people are often unaware of their own unawareness (Nisbett & Wilson, 1977).

Sometimes there are external sources of bias on people's reasons, such as cases in which people ask for only one type of reason (e.g., "Why did you like the movie?") or cases in which one type of reason is arbitrarily highly accessible in memory (e.g., because of people's recent experiences with the attitude object). Even in these cases, in which the source of bias on people's reasons is potentially easier to identify, people have difficulty recognizing that their reasons are biased. Consider the case in which reasons are biased by the way "why" questions are asked. Salancik (1974) and Seligman, Fazio, and Zanna (1980) manipulated the types of reasons people brought to mind by how they asked people for their reasons and found that people were insensitive to the fact that their reasons were biased. Seligman et al. (1980), for example, asked dating couples one of two types of questions about why they were attracted to their dating partner. Half of the couples were asked questions designed to elicit intrinsic reasons; for example, they were to complete the sentence "I date my girlfriend/boyfriend because I. . . ." The others were asked questions designed to elicit extrinsic reasons; for example, they were to complete the sentence "I date my girlfriend/boyfriend in order to. . . ." Interestingly, people did not seem to recognize that their reasons were biased by the wording of the questions, as indicated by the fact that they assumed that their reasons matched their current attitudes: Those in the intrinsic condition reported significantly more love for

their dating partner and reported that it was significantly more likely that they would marry their partner.

Seligman et al. (1980) and Salancik (1974) deliberately biased people's reasons in a rather subtle manner, making it difficult for people to recognize that their reasons were slanted in an internal or external direction. What about when there is a more obvious source of bias on people's reasons? For example, suppose that in the Greenwald (1968) study participants had been asked to think about why they felt the way they did about the issue of foreign aid rather than asked simply to recall the arguments. The act of generating reasons, we suggest, increases the perceived validity of thoughts that come to mind. Thoughts known to come from a biased, external source are transformed into reasons perceived to come from an internal source, thereby making these thoughts seem more credible and applicable to how one feels. Thus, we suspect that in our hypothetical reasons condition of the Greenwald (1968) study, people would generate reasons based on the arguments that were most accessible and, unlike those who simply recalled arguments, base their attitudes on these accessible thoughts.

More specifically, analyzing reasons can increase the perceived applicability of an accessible thought in at least three ways. First, when a thought comes to mind as part of a search for reasons, its biased source might be less evident than when people simply recall information. If people simply recall a particular argument against foreign aid, it is relatively easy to remember its biased source (the extra rehearsal of con arguments at the original session). If people are searching for reasons, however, and negative thoughts about foreign aid come to mind, the biased source of this information may be more difficult to recognize. A thought that comes to mind as a result of a search for reasons is, by definition, more diagnostic than a simple recollection; it is, after all, a reason for a feeling, and thus it must be at least somewhat diagnostic of that feeling.

Second, the act of analyzing reasons might cause people to go beyond the information recalled, generalizing from specific memories to more general qualities of the attitude object. In our studies, for example, we often find that when people give reasons about why they like or dislike another person, they do not simply list specific actions that person performed but describe that person's traits and qualities. A trait seems much more applicable to one's attitude toward another person than a memory of a specific act that might be biased or inapplicable. Finally, thinking about reasons might trigger memories that are consistent with accessible information (Ross, Lepper, & Hubbard, 1975). When reasons are analyzed, for example, the information that is accessible might remind people of specific memories that are perceived as highly relevant to their attitude.

In summary, we suggest that analyzing reasons will change people's attitudes only when two conditions are met. First, thoughts must be accessible in memory that are at least somewhat inconsistent with people's prior evaluation of the attitude object. Second, people must assume that these thoughts are applicable to their current attitude; that is, they must infer that these accessible thoughts reflect how they feel. This latter process, we predict, is more likely to occur when people analyze reasons than when they simply recall information about the attitude object, because thinking about reasons increases the per-

ceived applicability of accessible thoughts, even if those thoughts are known to come from a biased source.

We tested these hypotheses in Study 1 by manipulating the valence of the information that was accessible in people's memory about a target person after they had formed an initial impression of this person. Positive information was made more memorable for some participants (positive memory boost), whereas negative information was made more memorable for the other participants (negative memory boost). Next, after a delay, we asked people either to think about why they felt the way they did about the target person (the reasons condition) or to recall everything they could about her (the recall condition). All participants then rated how much they liked the target person. We predicted that people in the recall condition would not base their impressions on their accessible thoughts, as occurred in the study by Greenwald (1968). People in the reasons condition should be more likely to infer that their accessible thoughts reflect their attitude. If so, their impression should be more influenced by the memory manipulation (i.e., it should be more positive in the positive boost than the negative boost condition).

Study 1

Method

Participants

Fifty female students enrolled in introductory psychology participated in return for course credit. Women only were used because of possible sex differences in the kinds of attributes people like in others; previous research on the effects of analyzing reasons has not found systematic sex differences (e.g., Wilson & Schooler, 1991). We eliminated 2 participants from the analyses, 1 because she guessed the purpose of the study and 1 because she fell asleep during the study.

Procedure

Participants, seen individually or in groups of 2 to 5, were given both written and verbal instructions indicating that the study was concerned with impression formation. They were told that they would receive information about and be asked to form an impression of another student. The experimenter explained that participants would be given information about the student in two formats, ostensibly to examine how people use different kinds of information to form impressions. He handed out a list of 14 descriptions of the target person and asked participants to form an initial impression of her. Participants were given 3 min to read over the descriptions, after which the experimenter collected the lists. In pretesting, 5 of the descriptions had been rated as moderately positive (e.g., "Fran lent money to an acquaintance who had lost his wallet"), 5 as moderately negative (e.g., "Around exam time and other times when she is anxious, Fran is quite high strung"), and 4 as neutral (e.g., "Fran's upbringing was much like any other normal American girl"). Participants were then given 5 min to complete a series of filler questions, such as whether they had corrected vision or spoke another language.

Memory manipulation. We attempted to boost people's memory for either positive or negative descriptions as follows: The experimenter told participants that they would see the same information about the target person in the second format, on slides, with one behavior listed per slide. He began to show slides for 10 s each. After five slides had been shown, however, the projector appeared to malfunction. The experimenter examined the projector carefully and, with some dismay, said the bulb had burned out. This malfunction was staged to manipulate which charac-

teristics would be boosted in participants' memory. In the positive memory boost condition, three of the five slides shown before the staged mishap were positive and two were neutral. In the negative memory boost condition, three of the five slides were negative and two were neutral.[1] The experimenter told participants he was going to look for another bulb and that they should fill out the personality measure that they found in their folder while he was looking. The personality measure served as a filler task, and people were given 17 min to complete it.

Reasons analysis manipulation. When the experimenter returned, he said that he could not find another bulb for the projector and that the graduate student in charge of the study had instructed him to simply continue with the study. People were instructed to complete the next questionnaire in their folder. In the reasons condition, this questionnaire asked people to analyze and list the reasons describing why they felt the way they did about the target person. The instructions stated that the purpose of this was to organize people's thoughts and that they would not be asked to hand in the questionnaire. In the recall condition, participants were instructed to recall as many of the target person's behaviors as they could on a sheet provided with 14 boxes; participants placed one behavior in each box. The experimenter was unaware of whether people were in the reasons or recall condition. People were given 10 min to complete the reasons or recall questionnaire.[2]

Dependent Measures

All participants then rated their liking for the target person on four 7-point scales ranging from *not at all* (1) to *very much* (7). The four scales assessed how much they liked the target, how good a housemate they thought she would make, how much they would like to work on a project with her, and how much they would want to spend free time with her. As a check of how plausible participants found the information about the target person, they also rated how likely they thought they were to encounter such a person on a 7-point scale ranging from *not at all likely* (1) to *very likely* (7).

A subsequent packet instructed people in the recall condition to return to their recall questionnaire and rate each of the things they remembered, according to how positive or negative they considered it to be toward the target person, on a 7-point scale ranging from *very negative* (1) to *very positive* (7). People in the reasons condition were instructed to go back to their list of reasons and partition what they had written into individual reasons. What constituted a reason and where the divisions were made were left up to each individual. After partitioning their reasons, people rated the positivity of each reason on the same 7-point scale used by people in the recall condition. In addition, a research assistant coded people's reasons according to which of the 14 behaviors of the target person they mentioned (e.g., "I like her because she is generous to give money to her friend who lost his wallet") or whether they did not specifically mention any one of the target's behaviors (e.g., "I think I like her because she is similar to my friend"). A second rater coded a subset of the reasons, and she agreed with the first rater 94% of the time. At the conclusion of the study, all participants were fully debriefed.

Results

Manipulation Checks

People appeared to find the target person believable. Across conditions, their mean rating on the 7-point scale of how likely they thought they were to encounter a person like the target was 6.08 (*SD* = 0.96). These ratings did not vary by condition, $Fs(1, 44) < 2.25, ps > .14$.

The memory manipulation successfully boosted recall for either positive or negative behaviors. People in the recall condition recalled 74.6% of the behaviors that were repeated on slides, whereas they recalled only 37.4% of the behaviors not repeated, $t(21) = 8.61, p < .001$. People in the reasons condition mentioned in their reasons 35.4% of the behaviors that were repeated, whereas they mentioned only 19.2% of the behaviors that were not repeated, $t(25) = 3.67, p < .001$. As an additional manipulation check, we tallied the number of positive and negative characteristics of the target person that people mentioned in their recall or reasons, regardless of whether these characteristics were repeated on the slides. A 2 (memory boost: positive vs. negative) × 2 (condition: reasons vs. recall) × 2 (items recalled: positive vs. negative) between-subjects–within-subject analysis of variance (ANOVA) revealed the expected interaction between memory condition and number of positive versus negative characteristics recalled, $F(1, 44) = 18.06, p < .001$. People who saw positive slides listed more positive characteristics than negative characteristics in their recall or reasons (*Ms* = 2.32 and 1.16), whereas people who saw negative slides listed more negative characteristics than positive characteristics in their recall or reasons (*Ms* = 1.96 and 1.39).[3] Thus, the memory manipulation succeeded in biasing people's recall toward either positive or negative characteristics of the target person.

Liking for the Target Person

A composite measure of how much people liked the target person was computed by averaging people's responses to the

---

[1] The slides were always shown in the same order in which people had received the behaviors in written form. In the positive boost condition, the positive behaviors appeared in Positions 2, 4, 5, 8, and 12; neutral behaviors appeared in Positions 1, 3, 10, and 14; and negative behaviors appeared in Positions 6, 7, 9, 11, and 13. In the negative boost condition, the positions of the positive and negative behaviors were reversed. Thus, because the first five behaviors were shown again on slides, either positive or negative ones were repeated. The specific negative and positive behaviors that appeared in each position were varied in 1 of 20 random orders.

[2] It may appear that we introduced a confound by asking some people to analyze reasons and others to recall the target person's behavior. A more appropriate control might have been for people in the recall condition to complete no questionnaire rather than the recall task. We believe that giving people the recall task was better for two reasons. First, in previous studies we have compared people who analyzed reasons with a variety of control groups, including ones in which people completed filler tasks, ones in which people analyzed reasons about an unrelated topic, and ones in which people did nothing (e.g., Wilson & Schooler, 1991). The effect of analyzing reasons was the same in all cases and was similar to what we found here: People who analyzed reasons were more likely to change their attitudes than people who did not. Second, giving people the recall task controlled for merely thinking about the target person. Any differences found between the reasons and recall condition, we reasoned, could be attributed not merely to thinking about the target person but to something distinctive about analyzing reasons. It might be argued, however, that our design makes it unclear whether any differences are due to the reasons or the recall task. We address this issue later by reporting additional data comparing people who completed the recall task with a new control condition in which people did not complete any questionnaire.

[3] There was also a significant main effect of the reasons manipulation, $F(1, 44) = 14.95, p < .001$, reflecting the fact that people listed more items on the free-recall measure used in the recall condition (*M* = 4.45) than on the reasons questionnaire used in the reasons condition (*M* = 2.54). No other effects were significant.

four main dependent measures: how much they liked her, how good a housemate she would make, how much they would like to spend free time with her, and how much they would like to work on a project with her (Cronbach's α = .75). We predicted that the memory manipulation would influence people's liking in the reasons condition but not in the recall condition. This prediction was confirmed, as seen in Figure 1. A 2 (memory boost: positive vs. negative) $\times$ 2 (condition: reasons vs. recall) ANOVA revealed a significant interaction, $F(1, 44) = 4.22, p < .05$. The simple effect of the memory manipulation was highly significant in the reasons condition, $F(1, 44) = 7.76, p < .01$, but nonsignificant in the recall condition, $F(1, 44) < 1$.

Other Dependent Measures

One of our hypotheses was that people in the reasons condition would view their accessible thoughts as more applicable to their attitudes than people in the recall condition. To test this hypothesis, we correlated people's liking for the target person with the positivity of what they wrote in their reasons (in the reasons condition) or what they could recall about the target person (in the recall condition). In both conditions, the positivity scores were the mean of participants' own ratings of how positive their reasons or recollections were. As predicted, the correlations were higher in the reasons condition (rs = .86 and .71 in the negative and positive memory boost conditions) than

Liking for the Target Person

Figure 1. Liking for the target person: Study 1. The higher the rating, the greater the reported liking.

the recall conditions (rs = .14 and .16 in the negative and positive memory boost conditions). We performed a 2 (condition: reasons vs. recall) $\times$ 2 (memory boost: positive vs. negative) ANOVA on these correlations using techniques described by Games (1978), Snyder and Kendzierski (1982), and Steiner and Darroch (1969). The only significant finding was the predicted main effect of reasons condition, $z = 2.79, p < .01$.

We should note, however, that there is an alternative explanation of these results. The correlations were based on different measures in the two conditions: on people's free recall in the recall conditions and on people's reasons in the reasons conditions. As noted by Moser (1992), correlations based on recall are likely to be lower even if people base their judgments on their accessible thoughts, because people are likely to recall characteristics that were extraneous to their evaluations. In other words, a reasons measure might be better at assessing memory-based processing (Hastie & Park, 1986), because people might be more likely to list factors on this measure that are relevant to their judgment. The effects of the reasons manipulation on the liking ratings are consistent with our interpretation, in that people in the reasons condition were more influenced by the memory boost manipulation than people in the recall condition, suggesting that members of the former group did, in fact, base their attitudes more on their accessible thoughts than members of the latter group (see Figure 1). Nonetheless, the alternative explanation of the correlational results should be kept in mind.

Finally, our coding of the reasons allowed us to examine another one of our hypotheses, namely that people who analyze reasons go beyond the information that is accessible in memory, translating it into more global reasons. As mentioned earlier, we coded each reason as to whether it referred to one of the 14 descriptions people read about the target person or not. This coding was done very liberally; for example, if a person said that "Fran is very generous," we counted it as referring to the description "Fran lent money to an acquaintance who had lost his wallet." An example of a reason that did not refer to one of the descriptions is "I liked her because I can see a great deal of myself in her." Even with this liberal coding, only 62% of people's reasons referred to a specific description. We subsequently recoded the reasons with more stringent criteria; a reason was counted as a restatement of one of the descriptions only if the person listed the description without going beyond it in any way (the exact words did not have to be used; rephrasings were counted as well). The results of this coding, which had a similar reliability as our previous one, were as follows: 79% of people's reasons were generalizations from a description to a trait category, 9% were personal memories or associations, and only 7% were restatements of one of the categories (3% were coded as "other"). An example of a generalization involved one person who began by listing a specific description of the target person but then generalized to personality traits: "The first thing that comes to mind is the fact that she lent some money to an acquaintance of hers—this tells me that she is most likely a kind and giving person, two qualities which I look for in a friend." These results are consistent with our hypothesis that, when people generate reasons, they often go beyond the information given, transforming accessible thoughts into reasons that seem like valid indicators of one's attitude.
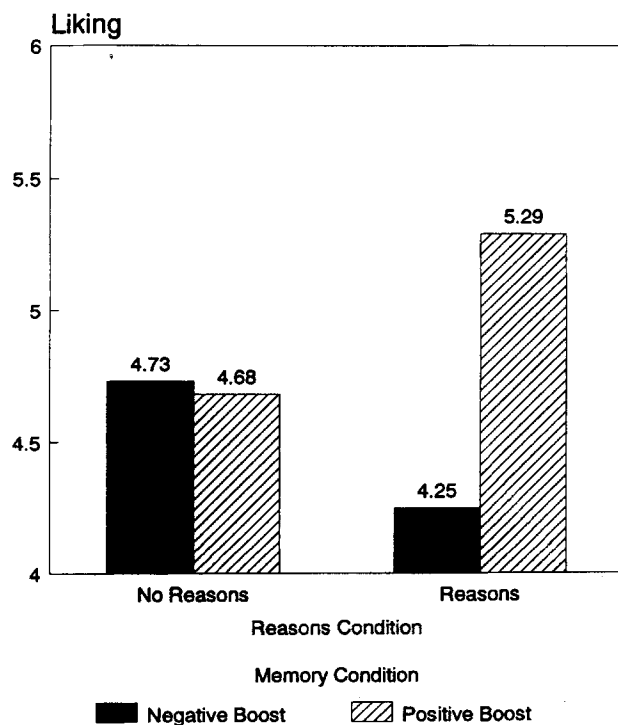
Follow-Up Data

Follow-up data were collected from two samples to aid in the interpretation of the results of the main study. We assumed that people in the recall condition were more likely to recognize that their accessible thoughts were biased by the slide projector malfunction and thus to rely more on their initial impressions of the target person. We were not able to assess this prediction directly in the main study because we did not ask people for their initial impressions (i.e., how they felt about the target person right after reading the initial descriptions of her). To remedy this, we conducted a follow-up with 78 new female participants. The procedure was identical to that of the main study up to the point when people read the 14 descriptions, at which time we gave them the dependent measures of how much they liked the target person. The order of the descriptions was the same as that used in the positive and negative memory boost conditions of the main study (see Footnote 1).

Consistent with our predictions, people's initial impressions in the follow-up were very similar to the impressions of people in the recall condition of the main study. When people received the descriptions in the negative order, their mean impression rating was 4.55; when they received the descriptions in the positive order, their mean impression rating was 4.74. These means did not differ significantly from each other, $t(76) = 1.04$, ns, which is consistent with our assumption that the initial order of the descriptions did not influence people's impressions. Furthermore, there was no significant difference between the ratings of the follow-up participants and people in the recall condition, which is consistent with our conclusion that people in the recall condition relied on their initial impressions. There was a slight indication that people in the recall condition overcorrected their impressions, in that those in the positive memory boost condition rated the target person more negatively than follow-up participants and those in the negative memory boost condition rated the target person more positively than follow-up participants. However, the interaction between memory boost and whether people were in the recall or follow-up conditions did not approach significance, $F(1, 120) < 1$.[4] The interaction between the memory boost and whether people were in the reasons or follow-up conditions was significant, $F(1, 120) = 3.91$, $p = .05$, indicating that, as expected, the attitudes of people who analyzed reasons differed significantly from initial impressions based on just reading the descriptions of the target person.

As noted earlier (see Footnote 2), a critic might argue that there is a potential confound in the design of Study 1 in that people in the reasons condition thought about why they felt the way they did about the target person before rating their attitudes, whereas people in the recall condition recalled everything they could about her. It might be argued that a control condition should have been included in which people did not receive the reasons or recall task to determine whether analyzing reasons increased the judged usability of accessible thoughts, whether the recall task decreased the judged usability of accessible thoughts, or both. The results of the first follow-up are consistent with our interpretation that the reason manipulation changed people's attitudes, whereas asking people to recall information did not. Nonetheless, to address this question further, we conducted a second follow-up. This was an exact replication

of the recall condition in the main study, with the addition of a new condition in which people were not asked to recall anything before rating the target person. More specifically, we conducted a partial replication of Study 1 with 54 new participants. For half of the participants, we replicated the recall condition exactly, in which people were asked to recall as many of the target person's behaviors as they could before rating their attitude toward her. We included a new condition that was identical, except that people were not asked to recall the target person's behaviors before they rated her. We also included the identical memory boost manipulation designed to get people to remember better positive or negative behaviors.

As in the main study, the memory boost manipulation was very successful. The interaction between the positive versus negative memory boost and the number of positive versus negative characteristics recalled was highly significant, $F(1, 50) = 84.68$, $p < .0001$, reflecting the fact that people who received the positive boost recalled more positive than negative characteristics of the target person ($Ms = 2.57$ and 1.11), whereas those who received the negative boost recalled more negative than positive information ($Ms = 2.42$ and 1.23).[5]

As predicted, the memory manipulation had no detectable effect on people's attitudes toward the target person, $F(1, 50) < 1$, ns. Furthermore, the memory manipulation was similarly ineffective in both the recall condition (the replication of Study 1) and the no recall condition. In the recall condition the mean attitude ratings, in the positive and negative recall conditions were 4.35 and 4.43, respectively, whereas in the no recall conditions the means were 4.81 and 4.79, respectively. The interaction between recall and memory boost was not significant, $F(1, 50) < 1$, ns. Although people in the recall condition rated the target person somewhat lower than people in the no recall condition, the main effect of recall was not significant, $F(1, 50) = 2.26$, $p = .09$.

It is instructive to combine the data from the main study and both follow-ups, which together constitute a 4 (initial impressions from Follow-Up 1, no recall condition from Follow-Up 2,

---

[4] It is interesting to consider why contrast effects did not occur in the recall conditions. In the literature on priming, making people aware of the arbitrary nature of accessible thoughts often (but not always; see Higgins, in press) results in contrast effects, whereby people evaluate the stimulus in a direction away from the accessible thoughts. It thus might be expected that people in the recall condition, who were hypothesized to be most aware of the invalidity of their accessible thoughts about the target person, would overadjust their impressions, resulting in a contrast effect. Unlike in most priming studies, however, people in the present studies had formed a prior impression of the target person that they could fall back on, instead of trying to correct for the amount of perceived bias on their impressions. That is, they could simply recall how they felt earlier and report this impression rather than undergoing an adjustment that might result in overcorrection.

[5] The people in the no recall condition were given the recall questionnaire right after rating their attitudes, allowing us to test the effectiveness of the memory manipulation in both the recall and no recall conditions. The manipulation was effective in both conditions, as reflected by the fact that the interaction between positive versus negative memory boost and positive versus negative information recalled was highly significant in both the recall and no recall conditions, $Fs(1, 50) > 25.18$, $ps < .0001$.

combination of the recall conditions from the main study and Follow-Up 2, or reasons condition from the main study) $\times$ 2 (memory boost) design. (Because people did not receive the memory boost in Follow-Up 1, in this case the memory manipulation reflects whether the order of the descriptions was that used in the negative or positive boost condition.) We should note that because this design combines data across studies, people were not randomly assigned to each condition. Although the results should be interpreted with caution, it is nonetheless interesting that the memory boost manipulation influenced people's attitudes only in the reasons condition (see means in Figure 2). The Condition $\times$ Memory Boost manipulation was nearly significant, $F(3, 172) = 2.42, p = .07$. More focused contrasts revealed that the effect of the memory boost manipulation was significantly larger in the reasons condition than in any of the other three conditions, $Fs(1, 120) > 4.56, ps < .05$. For example, a contrast that assigned weights of 1, -1, -1, and 1 to the no recall–negative boost, no recall–positive boost, reasons–negative boost, and reasons–positive boost cells was significant, $F(1, 120) = 7.14, p < .01$.

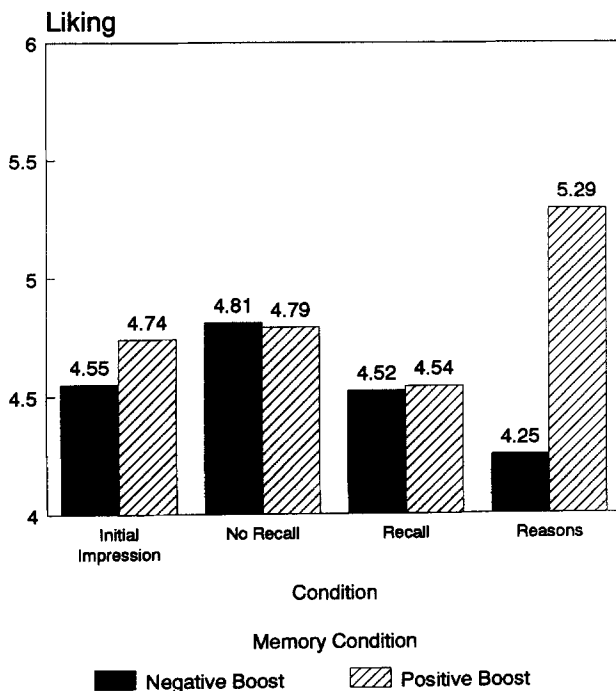## Liking for the Target Person: Main Study and Follow-Ups

Figure 2. Liking for the target person: Study 1 and follow-ups. The initial impression conditions are those in which people rated their liking for the target person right after reading the descriptions of her. In these conditions, the negative boost refers to those who received the descriptions in the same order as those in the negative boost conditions, and the positive boost refers to those who received the descriptions in the same order as those in the positive boost conditions. The higher the rating, the greater the reported liking.

Discussion

The results of Study 1 and the follow-ups support our hypothesis that the effects of analyzing reasons depend on the accessibility of what people recall about a stimulus. The memory manipulation succeeded in making positive thoughts about the target person more accessible in one condition and negative thoughts more accessible in another. When people analyzed reasons, they used these thoughts to explain why they felt the way they did and based their attitudes on these accessible thoughts. The positive versus negative thoughts were equally accessible in the recall conditions, but, interestingly, people in these conditions did not base their attitudes as much on their accessible thoughts. Instead, they appear to have relied on the impression they formed initially, disregarding the thoughts that came to mind when they recalled everything they could about the target person. This finding is a replication of the aforementioned study by Greenwald (1968), in which people's attitudes toward foreign aid were uninfluenced by a manipulation that made pro versus con arguments accessible in memory. As argued earlier, we suggest that people in Greenwald's study and in our recall conditions were likely to recognize that their accessible thoughts were biased, in Greenwald's case by the experimenter's instruction to rehearse only half of the arguments and in our case by the malfunction of the slide projector. In Higgins's (in press) terms, they judged their accessible thoughts to be low in "judged usability." When people analyzed reasons, they were, as predicted, more likely to infer that their accessible thoughts reflected their attitudes.

We do not mean to imply that people in the reasons conditions based their attitudes solely on what they could remember about the target person, totally ignoring the impressions they had formed earlier. Nor do we mean to suggest that people in the recall conditions based their attitude solely on their initial impressions, totally ignoring the behaviors they could remember. As noted by Carlston and Skowronski (1986), people's memories for their prior impressions and memories for specific behaviors of a target person can both be activated and combined to form a new impression. Our point is that people who analyzed reasons placed more weight on their accessible memories of the target person's behaviors than did people in the recall conditions.

An interesting question is what would happen, in the recall condition, if people did not have a prior impression to fall back on. If people in the recall condition had not formed an initial attitude of the target person, would they have relied more on their accessible thoughts? Would analyzing reasons increase a reliance on accessible thoughts even if people had no initial attitude? One possibility is that analyzing reasons will not make a difference under these circumstances because people, regardless of whether they think about reasons, will base their attitude on their accessible thoughts about the target person. We predict, however, that analyzing reasons will increase the extent to which people will base their attitudes on accessible thoughts, even when people have no initial attitude.

When people do not analyze reasons and have no initial attitude, they are still likely to recognize that their accessible thoughts are not completely valid (e.g., because of the malfunction of the slide projector) and thus are likely to avoid basing

their attitude too much on these thoughts. One way they might do so is by adjusting their impressions away from the direction of their accessible thoughts; for example, if their accessible thoughts are positive, they might adjust their impressions in a negative direction somewhat. When people with no initial attitude analyze reasons, they should be more likely to view their accessible thoughts as applicable and valid and thus should be more likely to base their attitudes on their accessible thoughts. They have no earlier impression to fall back on and are most likely to view their accessible thoughts as applicable to their attitude.

To test these hypotheses, we borrowed a manipulation of whether people form an initial attitude from the literature on on-line versus memory-based processing (Hastie & Park, 1986), in which people are either told or not told, before receiving information about a person, how they will be evaluating this person. In Study 2, we told half of the participants, before giving them information about the target person, that they would be asked to evaluate how good a social worker she would make. There participants were expected to form initial, on-line evaluations on this dimension. The others were not told that they would be asked the social worker question until just before rating the target person (i.e., after the memory manipulation). All participants received the positive memory boost. Thus, a 2 (judgment: on line vs. memory based) × 2 (condition: reasons vs. recall) design was used. We predicted that analyzing reasons would increase people's reliance on their accessible thoughts in the on-line condition, in which they had formed initial impressions of the target person. This would be a replication of Study 1. Analyzing reasons should also increase people's reliance on their accessible thoughts in the memory-based condition by increasing the perceived validity of these thoughts.

Study 2

Method

Participants

Participants were 82 female undergraduates. Some were introductory psychology students who received course credit for their participation, whereas others were summer school students who were paid for their participation. Two participants were eliminated from the analyses because they guessed the purpose of the study.

Procedure

The procedure of Study 2 was identical to that of Study 1, except for the following changes. Rather than asking people how much they liked the target person, the chief dependent measure was how good a social worker they thought she would be. We pretested behaviors for the target person that were relevant to this judgment: Five were viewed as evidence that the person would make a good social worker (e.g., "Fran is known to give objective advice"), five were viewed as evidence that the person would not make a good social worker (e.g., "Fran gets lonely and troubled when she spends time by herself"), and four were neutral with regard to how good a social worker she would be (e.g., "Fran voted for a politically moderate candidate in the last congressional election"). Instead of manipulating whether positive or negative behaviors were most memorable, however, positive behaviors were made more memorable for all participants. This was done in the same fashion as in Study

1, by repeating three of them on slides before the projector ostensibly malfunctioned.

Manipulation of on-line versus memory-based processing. We manipulated whether people's judgments of the target person's qualifications as a social worker were made in an on-line or memory-based fashion. Following Hastie and Park (1986), we did so by informing half of the participants in advance, before they read about the target person's behavior, that they would be asked to make this judgment. People in the memory-based condition were not told that they would be rating her ability as a social worker until immediately before they either analyzed reasons or recalled her behaviors.

Reasons manipulation. From this point on, the procedure was identical to that of Study 1. After the 17-min filler task, half of the participants listed, privately and anonymously, reasons why the target person would or would not make a good social worker. Again, they were told that the purpose of this task was to organize their thoughts and that they would not be asked to hand in the questionnaire. People in the recall condition were asked to recall all of the characteristics of the target person that they could.

Dependent Measures

All participants then rated, on 7-point scales, how good a social worker the target person would be (1 = *not very good,* 7 = *very good*), how effective she would be if she got a job as a social worker (1 = *not very effective,* 7 = *very effective*), how well suited she was for a job as a social worker (1 = *not well suited,* 7 = *very well suited*), and how strongly they would recommend that she enter the field of social work (1 = *not at all,* 7 = *very strongly*). As a check of how plausible they found the information about the target person, people also rated how likely they thought they were to encounter such a person on a 7-point scale ranging from *not at all likely* (1) to *very likely* (7).

Results and Discussion

Once again, most participants found the target person to be quite believable. The mean rating on the 7-point scale of how likely they thought they would encounter such a person was 5.73 (SD = 0.99). In addition, the memory manipulation was successful. In the recall condition, people recalled 80% of the behaviors that were repeated but only 32% of the behaviors that were not, $t(38)$ = 12.56, p < .001. In the reasons condition, people mentioned in their reasons 48% of the behaviors that were repeated and only 18% of the behaviors that were not, $t(40)$ = 7.48, p < .001.

We averaged participants' four ratings of how good a social worker the target person would be (Cronbach's α = .94). A 2 (judgment: on line vs. memory based) × 2 (condition: reasons vs. recall) ANOVA on this index revealed the predicted two main effects (see top of Figure 3). People in the memory-based condition made more positive ratings than people in the on-line condition, $F(1, 76)$ = 4.39, p < .04, which is consistent with the prediction that they would base their ratings more on what they could remember. This confirms the effectiveness of the memory-based versus on-line manipulation. Second, people in the reasons condition made more positive ratings than people in the recall condition, $F(1, 76)$ = 4.20, p < .05. The interaction was not significant, $F(1, 76)$ < 1. Thus, as predicted, analyzing reasons increased the extent to which people based their impressions on their accessible thoughts, independently of the extent to which people's initial attitude was memory based.

To test this prediction further, we computed the correlation

Suitability as a Social Worker

Liking
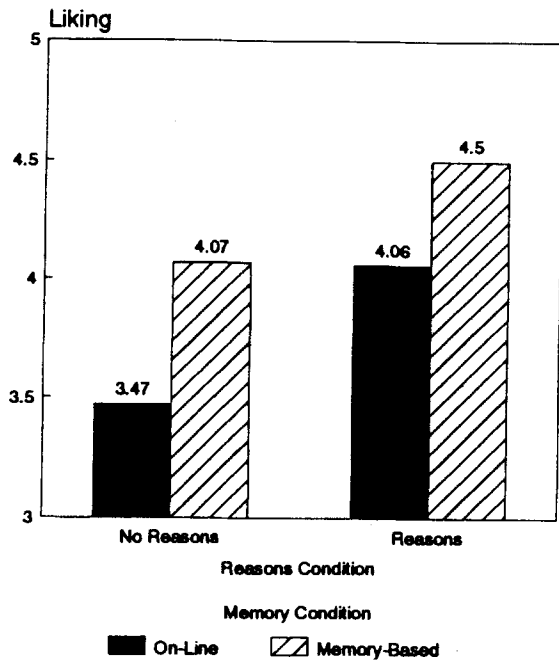
Correlation Between Suitability Ratings and Recall/Reasons
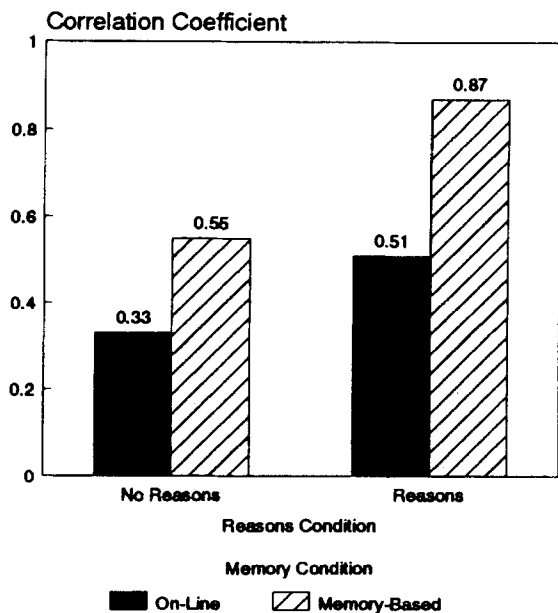
Correlation Coefficient

Figure 3. Ratings of how good a social worker the target person would be (top panel) and the correlation between these ratings and recall (bottom panel): Study 2. In the top panel, the higher the rating, the more people reported that the target person would make a good social worker.

between people's evaluations of how good a social worker the person would make and their accessible thoughts. In the recall condition, in which people recalled the target's attributes before rating her, the measure was the correlation between people's average ratings of the positivity of these recollections and their evaluation of the target. In the reasons condition, in which people listed reasons before rating her, the measure was the correlation between people's average rating of the positivity of their reasons and their evaluation of the target. As seen at the bottom of Figure 3, these correlations were significantly higher in the memory-based condition, $z = 2.14$, $p = .03$, and higher in the reasons condition, $z = 1.94$, $p = .05$. The interaction was not significant, $z = 1.04$, ns. These results are consistent with our prediction that people would base their evaluations more on their accessible thoughts when they analyzed reasons and when they had not formed initial, on-line impressions. As in Study 1, the fact that the correlations were based on different measures in the reasons and recall conditions should be kept in mind, however, because it is possible that the reasons measure is better able to assess memory-based processing than a recall measure (Moser, 1992). At a minimum, the correlations serve as an additional check on the manipulation of on-line versus memory-based processing. The fact that the correlations were significantly higher in the memory-based versus the on-line conditions, independently of whether they were based on recall or reasons, is further evidence that the manipulation was successful.

Combined with the predicted effects on evaluations shown in the top half of Figure 3, the correlational results lend credence to our conclusion that analyzing reasons can change attitudes both when people's initial attitude is relatively on line and when it is more memory based. In other words, analyzing reasons increases the extent to which people base their attitudes on accessible thoughts about the attitude object, even when people do not have an initial attitude to fall back on (in the memory-based condition). When people did not analyze reasons in the memory-based condition, they appear to have moderated their judgments some, presumably because they recognized that their accessible thoughts were biased by the malfunction of the slide projector. When people did analyze reasons they showed less moderation, presumably because they viewed their accessible thoughts as more applicable to their attitudes.

General Discussion

We found, in two studies, that asking people to think about why they felt the way they did about a target person changed their attitudes in the direction of the valence of thoughts that were accessible in memory. When positive thoughts about the target person were accessible in memory, as a result of our memory manipulation, people who analyzed reasons changed their attitude in a positive direction. When negative thoughts about the target person were accessible in memory, people who analyzed reasons changed their attitude in a negative direction. People in the recall conditions seemed to recognize more that their accessible thoughts were not entirely applicable to how they felt and thus did not base their attitudes as much on these thoughts.

Are there conditions under which people who analyze reasons will recognize that their reasons are biased, preventing them-

selves from inferring that their reasons reflect their attitudes? As mentioned earlier, we have found one such condition: when people's initial attitude is highly accessible (Hodges & Wilson, 1994). What about when people's initial attitude is not highly accessible, as in the present studies? Perhaps if the source of bias on people's reasons was even more obvious than in the present studies, people would be more likely to prevent themselves from inferring that their reasons reflected how they felt. For example, it is interesting to consider what would happen if people were directly forewarned about the lack of validity of their reasons. If people were told that it is difficult to verbalize one's reasons, perhaps they would be more reluctant to infer that their reasons match their attitudes.

We suspect, however, that such a forewarning might not be effective and, in fact, might even backfire. We have argued that the process of generating reasons increases the perceived validity of the thoughts that come to mind. Telling people that their reasons might be biased is unlikely to prevent this process from occurring. Even when forewarned, people are likely to go beyond the information contained in their accessible thoughts, generating memories and associations that are consistent with these thoughts. Forewarning people might even exacerbate this process. When cautioned about the possible invalidity of their reasons, people might exert more of an effort to come up with unbiased reasons. Ironically, this might result in even more of an attempt to go beyond the information in one's accessible thoughts, generating even more memories and associations that are consistent with these thoughts. As discussed earlier, the process of going beyond specific memories about the attitude object is likely to increase the perceived applicability of one's reasons. Thus, if forewarning people causes them to be even more likely to generalize from their accessible thoughts, they might be even more likely to assume that their reasons reflect their current attitudes.

Wilson and Erber (1994) obtained evidence that is consistent with this prediction. They conducted a replication of a study by Wilson et al. (1993) in which people evaluated art posters of the sort that college students often purchase and put on their walls. As in the Wilson et al. (1993) study, people evaluated two types of posters: two reproductions of classic paintings (one by Van Gogh and one by Monet) and three contemporary, humorous posters (e.g., one was a photograph of a kitten perched on a rope with the caption "Gimme a Break"). Some participants were asked to write down, privately and anonymously, why they felt the way they did about each poster, whereas others completed a filler task. All participants then rated their liking for each poster and indicated which one they would choose if they could take one home.

We added a third condition that was identical to the reasons condition, except that people were forewarned that the reasons people generate are not always applicable to how they feel. Before people introspected about reasons, they received the following instructions: "People often do not really know why they like or dislike something. Even if people know, they often have a hard time putting their reasons into words. So, this questionnaire might not be very easy to fill out. Just do the best you can." Our goal was to communicate to people that the reasons they generated might not be a complete reflection of how they feel; that is, their reasons might not be applicable to their attitudes.

As in the Wilson et al. (1993) study, people who analyzed reasons (without the warning) were more likely to prefer the humorous to the art posters and were slightly more likely to choose one of the humorous posters to take home (these results were not as strong as in the original study but were in the same direction). As predicted, people who analyzed reasons and were forewarned were more likely than those in the standard reasons condition to change their attitudes toward the posters; they were significantly more likely to say they would choose one of the humorous posters to take home than people who analyzed reasons without the warning.

As expected, the forewarning manipulation increased the judged applicability of people's reasons rather than convincing people that their reasons were inapplicable. On a manipulation check question, we asked people how likely it was that there were important influences on their reactions to the poster that they could not put into words. People who received the warning said that there were significantly fewer influences that they could not put into words; that is, they believed that their reasons were more complete, and hence reflective of their attitudes, than people who analyzed reasons without the forewarning manipulation. Furthermore, there was evidence that people in the forewarning condition were more likely to generalize from their accessible thoughts about the posters. In comparison with those in the standard reasons condition, forewarned participants were less likely to list specific attributes of the posters and more likely to describe personal memories or affect triggered by the posters.

The Wilson and Erber (1994) results address a possible limitation of the present studies, namely that we did not directly assess how applicable or "usable" people believed their accessible thoughts to be. Although the results on the main dependent measures were very consistent with our prediction that analyzing reasons increases the judged usability of one's accessible thoughts, we did not include a measure that asked people directly how applicable they considered their accessible thoughts to be. Wilson and Erber (1994) did include such a measure, and, ironically, people who were forewarned were even more likely than those not forewarned to assume that their reasons were complete and unbiased.

In summary, the present results, along with Wilson and Erber's (1994), provide the first direct evidence for our hypotheses about the role of accessibility and judged usability in the effects of analyzing reasons. These hypotheses, we should note, are compatible with a conceptualization by Millar and Tesser (1986, 1989; Tesser, Martin, & Mendolia, in press). These authors focused on the distinction between the affective and cognitive components of an attitude and suggested that thinking about reasons increases the accessibility of cognition at the expense of affect. For example, a person's attitude toward a difficult crossword puzzle might initially be based on the affect it triggers (e.g., "I feel stupid when I can't solve it"). When people analyze reasons, however, they tend to focus on their cognitions about the attributes of the attitude object (e.g., "This puzzle might help my verbal skills"). These cognitions are now more accessible than the initial affect, leading to attitude change (e.g., to the extent that the cognitions are more positive than the affect, people have a more positive attitude).

The Millar and Tesser position is similar to our accessibility hypothesis, except that it argues that, after reasons have been analyzed, cognitions typically become more accessible and affect typically becomes less accessible. Millar and Tesser (1986, 1989) have obtained convincing evidence that such a process can occur, and we believe it is one way in which analyzing reasons can change attitudes. Our position, however, is that it is not the only way. We believe that analyzing reasons will change attitudes whenever the reasons that come to mind are of a different valence than people's initial attitudes, regardless of whether the newly accessible thoughts are of a cognitive or affective nature.

In the present studies, for example, there is no reason to assume that people's initial attitudes, formed when they first read the 14 behaviors of the target person, were more affective in nature than the attitudes people reported after analyzing reasons. Instead, people's initial attitudes were formed by combining information about all 14 behaviors, whereas people's later, reasons-based attitudes stemmed more from the subset of these behaviors that people could remember. The key, again, is whether the thoughts that are accessible when people analyze reasons match the valence of their initial attitudes and whether they infer that these thoughts reflect their current attitudes. Millar and Tesser's (1986, 1989) results are a special case of this process, in which people focus on cognitions that have a different valence than their initial affect. Our results suggest, however, that analyzing reasons will also change people's attitudes when the initial attitude is based on one set of cognitions (information about the 14 behaviors) and thinking about reasons focuses attention on a subset of these cognitions (the behaviors that are most memorable). Presumably, attitude change will also occur if people's initial attitude is affectively based, but analyzing reasons focuses people's attention on affect of a different valence.

We should note that Millar and Tesser (1986, 1989) have broadened our model in another respect: They have demonstrated that the extent to which attitudes will predict behavior depends on whether there is a match between the basis of the attitude (affective or cognitive) and the basis of the behavior (affective or cognitive) and that analyzing reasons increases the likelihood that an attitude will be cognitively based. This adds an intriguing specificity to the conditions under which people's attitudes are consistent with their behavior after introspections about reasons.

An alternative account of our results is that people in the reasons conditions perceived a demand to report an attitude that was consistent with the reasons they had just listed. Perhaps they were concerned about appearing inconsistent to the experimenter, making them reluctant to list reasons that implied one attitude and then reporting another attitude on a subsequent questionnaire. We believe, however, that a demand characteristics interpretation is deficient in several respects. First, we assured people that their reasons were private and anonymous and, in fact, would never be read by anyone. To the extent that people believed these instructions, they should have been unconcerned with appearing consistent to others. In addition, in a previous study, people thought about reasons but never wrote them down, and similar effects on attitude change were found (Wilson et al., 1984, Study 2). Finally, it is not clear why a demand to be consistent would not also be present in the condi-

tions of the present studies in which people were asked, just before reporting their attitude, to recall everything they could about the attitude object. To the extent that there was a demand to be consistent, people might have believed that their attitude should reflect what they could remember. In fact, we found that people's attitudes were not influenced by what they could recall in the recall conditions.

More generally, the present results speak to a current debate about how people form attitudes. On the one hand, several studies suggest that people's attitudes are a function of their prior evaluations, which are retrieved with no attempt to "recompute" the attitudes. For example, Bargh, Chaiken, Govender, and Pratto (1992) and Fazio, Sanbonmatsu, Powell, and Kardes (1986) found that people's prior attitudes are often retrieved automatically and without effort when the name of the attitude object is presented. There is disagreement as to how accessible an attitude has to be before it is activated automatically, but these theorists agree that people's attitudes are often a function of their prior evaluation, which is triggered when the attitude object is encountered (Chaiken & Bargh, 1993; Fazio, 1993).

In contrast, several researchers argue that people construct attitudes based on information that happens to be accessible at the time they are asked how they feel (e.g., Schwarz & Bless, 1992; Tourangeau & Rasinski, 1988; Zanna & Rempel, 1988). People's attitudes have been found to be altered by a number of manipulations of the thoughts that are accessible to them (e.g., Bem, 1972; Forgas, 1992; Schwarz & Clore, 1983; Tesser, 1978). For example, Bem's (1972) self-perception theory can be viewed as a constructivist theory of attitudes: Rather than relying on previously formed evaluations, people construct an attitude from observations of their behavior and the social context.

On the face of it, the retrieved evaluation and constructivist views of attitudes are contradictory: Why would people construct a new attitude if their prior evaluation is easily retrieved from memory or even activated automatically? It is becoming increasingly clear, however, that these views are not incompatible. There are moderator variables that determine when people rely on an evaluation stored in memory rather than construct an attitude. One such moderator is the strength of people's initial evaluation (Fazio, in press; Petty, Priester, & Wegener, 1994). If people's earlier evaluation of the attitude object is very accessible in memory, there should be less of a need to recompute the attitude on the basis of other information in memory.

The present studies point to another moderator variable: the extent to which people analyze the reasons for their attitudes. People who did not think about reasons were likely to rely on their previously formed evaluation of the target person, disregarding thoughts about the target person that were inconsistent with this evaluation. Those who analyzed reasons, however, were likely to recompute their attitude, basing it on accessible thoughts about the target person. Thus, even when people had formed an initial impression, those who analyzed reasons constructed a new attitude based on accessible information.

This process of recomputing an attitude raises some intriguing questions. There are at least two ways in which this construction process could take place: The newly constructed attitude could replace the prior one, or the newly constructed attitude could override, but not replace, the prior attitude. The

latter possibility would mean that people had a dual representation of the attitude object: the original evaluation plus the newly constructed one. There is evidence from the memory literature consistent with the view that initial judgments can be overridden instead of erased. Schooler and his colleagues have found that asking people to verbalize their memories for stimuli (e.g., faces) creates a new memory representation that "overshadows" but does not replace people's previous memories (Schooler & Engstler-Schooler, 1990; Schooler, Ohlsson, & Brooks, in press). For example, Schooler and Engstler-Schooler (1990) found that asking people to describe a face impaired later recognition of that face. However, when people were asked to describe the face and then were given only 5 s to perform the recognition test, their memory for the face was not impaired. Schooler and Engstler-Schooler (1990) concluded that verbalization does not erase the original memory for the face but "overshadows" the original memory with a new representation, which, under some conditions, impairs recognition performance.

Similarly, it is possible that verbalizing the reasons for one's attitude causes people to form a new evaluation that overrides the earlier one. Perhaps instructions similar to those used by Schooler and Engstler-Schooler (1990), whereby people are asked to report how they feel about the stimulus very quickly, would access the initial evaluation. Alternatively, there may be no initial attitude to access if it has been replaced by the newly constructed attitude. According to this view, there is only one attitude at any given time, but this attitude can change quickly depending on the information that happens to be accessible (assuming, that is, that a prior evaluation is not highly accessible).

These important questions about attitude representation and change are as of yet unsettled. It is clear, however, that the attitude change resulting from analyzing reasons can be consequential. If people make important choices based on their new, reasons-based attitude, they might regret that choice, to the extent that their chronic construction of the attitude object returns. Suppose, for example, that someone analyzes his or her reasons before purchasing a car and decides to buy the one that comes out most favorably in the reasons analysis. If the reasons that happened to be accessible imply a different preference than the person initially had about the cars and the initial attitude returns over time, he or she will regret the choice based on reasons. We have found evidence in a recent study for just this sequence of events (Wilson et al., 1993).

Fortunately, there are conditions under which thinking about attitudes will not change people's attitudes. When people are knowledgeable about the attitude object, they are less likely to bring to mind reasons that conflict with their initial attitude (Wilson, Kraft, & Dunn, 1989; Wilson et al., 1993). When an initial attitude is highly accessible, in that it can be brought to mind quickly, people are less likely to infer that their attitude matches the reasons that happen to be accessible (Hodges & Wilson, 1994). Thus, we do not mean to imply that every time people reach for a piece of paper and write down why they feel the way they do, they will construct a new attitude. Nonetheless, processes similar to those in the present studies occur not infrequently in everyday life: People are not very knowledgeable about the attitude object, they think about why they feel the way

they do, and the thoughts that are accessible imply a different attitude than they had before. It is under these conditions that people might want to hesitate in acting too readily on their reasons-based attitudes.

References

Bargh, J. A., Chaiken, S., Govender, R., & Pratto, F. (1992). The generality of the automatic activation effect. *Journal of Personality and Social Psychology, 62,* 893–912.

Bem, D. J. (1972). Self-perception theory. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 6, pp. 1–62). New York: Academic Press.

Carlston, D. E., & Skowronski, J. J. (1986). Trait memory and behavior memory: The effect of alternative pathways on impression judgment response times. *Journal of Personality and Social Psychology, 50,* 5–13.

Chaiken, S., & Bargh, J. A. (1993). Occurrence versus moderation of the automatic attitude activation effect: Reply to Fazio. *Journal of Personality and Social Psychology, 64,* 759–765.

Fazio, R. H. (1993). Variability in the likelihood of automatic attitude activation: Data reanalysis and commentary on Bargh, Chaiken, Govender, and Pratto (1992). *Journal of Personality and Social Psychology, 64,* 753–758.

Fazio, R. H. (in press). Attitudes as object-evaluation associations: Determinants, consequences, and correlates of attitude accessibility. In R. Petty & J. Krosnick (Eds.), *Attitude strength: Antecedents and consequences.* Hillsdale, NJ: Erlbaum.

Fazio, R. H., Sanbonmatsu, D. M., Powell, M. C., & Kardes, F. R. (1986). On the automatic activation of attitudes. *Journal of Personality and Social Psychology, 50,* 229–238.

Forgas, J. P. (1992). Affect in social judgments and decisions: A multiprocess model. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 25, pp. 227–275). New York: Academic Press.

Games, P. A. (1978). A three-factor model encompassing many possible statistical tests on independent groups. *Psychological Bulletin, 85,* 168–182.

Greenwald, A. G. (1968). Cognitive learning, cognitive response to persuasion, and attitude change. In A. G. Greenwald, T. C. Brock, & T. M. Ostrom (Eds.), *Psychological foundation of attitudes* (pp. 147–170). New York: Academic Press.

Hastie, R., & Park, B. (1986). The relationship between memory and judgment depends on whether the judgment task is memory-based or on-line. *Psychological Review, 93,* 258–268.

Higgins, E. T. (1989). Knowledge accessibility and activation: Subjectivity and suffering from unconscious sources. In J. S. Uleman & J. A. Bargh (Eds.), *Unintended thought* (pp. 75–123). New York: Guilford Press.

Higgins, E. T. (in press). Knowledge activation: Accessibility, applicability, and salience. In E. T. Higgins & A. R. Kruglanski (Eds.), *Social psychology: Handbook of basic principles.* New York: Guilford Press.

Hodges, S. D., & Wilson, T. D. (1994). Effects of analyzing reasons on attitude change: The moderating role of attitude accessibility. *Social Cognition, 11,* 353–366.

Johnson, R. W., MacArthur, A., & Wright, E. F. (1991). Timing of reasons analysis and attitude-behavior consistency. *Psychological Reports, 69,* 603–608.

Kruglanski, A. W., Friedland, N., & Farkash, E. (1984). Lay persons' sensitivity to statistical information: The case of high perceived applicability. *Journal of Personality and Social Psychology, 46,* 503–518.

Martin, L. L., & Achee, J. W. (1992). Beyond accessibility: The role of processing objectives in judgment. In A. Tesser & L. Martin (Eds.), *The construction of social judgment* (pp. 195–216). Hillsdale, NJ: Erlbaum.

Millar, M. G., & Tesser, A. (1986). Effects of affective and cognitive focus on the attitude-behavior relationship. *Journal of Personality and Social Psychology, 51,* 270-276.

Millar, M. G., & Tesser, A. (1989). The effects of affective-cognitive consistency and thought on the attitude-behavior relation. *Journal of Experimental Social Psychology, 25,* 189-202.

Moser, D. V. (1992). Does memory affect judgment? Self-generated versus recall memory measures. *Journal of Personality and Social Psychology, 62,* 555-563.

Nisbett, R. E., & Wilson, T. D. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review, 84,* 231-259.

Petty, R. E., Priester, J. R., & Wegener, D. T. (1994). Cognitive processes in attitude change. In R. S. Wyer & T. K. Srull (Eds.), *Handbook of social cognition* (Vol. 2, 2nd ed., pp. 69-142). Hillsdale, NJ: Erlbaum.

Ross, L., Lepper, M. R., & Hubbard, M. (1975). Perseverance in self perception and social perception: Biased attributional processes in the debriefing paradigm. *Journal of Personality and Social Psychology, 32,* 880-892.

Salancik, J. R. (1974). Inference of one's attitude from behavior recalled under linguistically manipulated cognitive sets. *Journal of Experimental Social Psychology, 10,* 415-427.

Schooler, J. W., & Engstler-Schooler, T. Y. (1990). Verbal overshadowing of visual memories: Some things are better left unsaid. *Cognitive Psychology, 22,* 36-71.

Schooler, J. W., Ohlsson, S., & Brooks, K. (in press). Thoughts beyond words: When language overshadows insight. *Journal of Experimental Psychology: General.*

Schwarz, N., & Bless, H. (1992). Constructing reality and its alternatives: An inclusion/exclusion model of assimilation and contrast effects in social judgment. In A. Tesser & L. Martin (Eds.), *The construction of social judgment* (pp. 217-245). Hillsdale, NJ: Erlbaum.

Schwarz, N., & Clore, G. L. (1983). Mood, misattribution, and judgment of well-being: Informative and directive functions of affective states. *Journal of Personality and Social Psychology, 45,* 513-523.

Seligman, C., Fazio, R. H., & Zanna, M. P. (1980). Effects of salience of extrinsic rewards on liking and loving. *Journal of Personality and Social Psychology, 38,* 453-460.

Snyder, M., & Kendzierski, D. (1982). Acting on one's attitudes: Procedures for linking attitude and behavior. *Journal of Experimental Social Psychology, 18,* 165-183.

Steiner, I. D., & Darroch, R. K. (1969). Relationship between the quality of counterattitudinal performance and attitude change. *Journal of Personality and Social Psychology, 11,* 312-320.

Strack, F. (1992). The different routes to social judgments: Experiential versus informational strategies. In A. Tesser & L. Martin (Eds.), *The construction of social judgment* (pp. 249-276). Hillsdale, NJ: Erlbaum.

Tesser, A. (1978). Self-generated attitude change. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 11, pp. 289-338). New York: Academic Press.

Tesser, A., Martin, L., & Mendolia, M. (in press). The role of thought in changing attitude strength. In R. Petty & J. Krosnick (Eds.), *Attitude strength: Antecedents and consequences.* Hillsdale, NJ: Erlbaum.

Tourangeau, R., & Rasinski, K. A. (1988). Cognitive processes underlying context effects in attitude measurement. *Psychological Bulletin, 103,* 299-314.

Wilson, T. D. (1990). Self-persuasion via self-reflection. In J. Olson & M. P. Zanna (Eds.), *Self-inference processes: The Ontario Symposium* (Vol. 6, pp. 43-67). Hillsdale, NJ: Erlbaum.

Wilson, T. D., & Brekke, N. C. (1994). Mental contamination and mental correction: Unwanted influences on judgments and evaluations. *Psychological Bulletin, 116,* 117-142.

Wilson, T. D., Bybee, J. A., Dunn, D. S., Hyman, D. B., & Rotondo, J. A. (1984). Effects of analyzing reasons on attitude-behavior consistency. *Journal of Personality and Social Psychology, 47,* 5-16.

Wilson, T. D., & Dunn, D. S. (1986). Effects of introspection on attitude-behavior consistency: Analyzing reasons versus focusing on feelings. *Journal of Personality and Social Psychology, 22,* 249-263.

Wilson, T. D., Dunn, D. S., Kraft, D., & Lisle, D. J. (1989). Introspection, attitude change, and attitude-behavior consistency: The disruptive effects of explaining why we feel the way we do. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 19, pp. 123-205). Orlando, FL: Academic Press.

Wilson, T. D., & Erber, M. W. (1994). [Effects of warning people about the validity of their reasons]. Unpublished raw data.

Wilson, T. D., & Hodges, S. D. (1992). Attitudes as temporary constructions. In A. Tesser & L. Martin (Eds.), *The construction of social judgment* (pp. 37-65). Hillsdale, NJ: Erlbaum.

Wilson, T. D., & Kraft, D. (1993). Why do I love thee?: Effects of repeated introspections about a dating relationship on attitudes toward the relationship. *Personality and Social Psychology Bulletin, 19,* 409-418.

Wilson, T. D., Kraft, D., & Dunn, D. S. (1989). The disruptive effects of explaining attitudes: The moderating effect of knowledge about the attitude object. *Journal of Experimental Social Psychology, 25,* 379-400.

Wilson, T. D., Lisle, D., Schooler, J., Hodges, S. D., Klaaren, K. J., & LaFleur, S. J. (1993). Introspecting about reasons can reduce postchoice satisfaction. *Personality and Social Psychology Bulletin, 19,* 331-339.

Wilson, T. D., & Schooler, J. (1991). Thinking too much: Introspection can reduce the quality of preferences and decisions. *Journal of Personality and Social Psychology, 60,* 181-192.

Wilson, T. D., & Stone, J. I. (1985). More on telling more than we can know. In P. Shaver (Ed.), *Review of personality and social psychology* (Vol. 6, pp. 167-183). Beverly Hills, CA: Sage.

Zanna, M. P., & Rempel, J. K. (1988). Attitudes: A new look at an old concept. In D. Bar-Tal & A. W. Kruglanski (Eds.), *The social psychology of knowledge* (pp. 315-334). Cambridge, England: Cambridge University Press.