Journal of Personality and Social Psychology
November 1988 Vol. 55, No. 5, 685-694

Of Thoughts Unspoken
Social Inference and the Self-Regulation of Behavior

Daniel T. Gilbert
University of Texas at Austin
Douglas S. Krull
University of Texas at Austin
Brett W. Pelham
University of Texas at Austin

ABSTRACT

How do attempts to regulate one's own behavior affect the inferences one draws about others? We suggest that perceivers draw dispositional inferences about targets (characterization) and then adjust those inferences with information about the constraints on the targets' behaviors (correction). Because correction is more effortful than characterization, perceivers who devote cognitive resources to the regulation of their own behavior should be able to characterize targets but unable to correct those characterizations. In Experiment 1, unregulated subjects incidentally ignored an irrelevant stimulus while they observed a target's behavior, whereas self-regulated subjects purposefully ignored the same irrelevant stimulus. In Experiment 2, unregulated subjects expressed their sincere affection toward a target, whereas self-regulated subjects expressed false affection. In both experiments, self-regulated subjects were less likely than unregulated subjects to correct their characterizations of the target. The results suggest that social interaction (which generally requires the self-regulation of ongoing behavior) may profoundly affect the way in which active perceivers process information about others.

Lying is not only excusable; it is not only innocent, and instinctive; it is, above all, necessary and unavoidable. Without the ameliorations that it offers life would become a mere syllogism, and hence too metallic to be borne. (H. L. Mencken, 1924 , p. 277)

What would social life be like if people invariably said what they meant and meant what they said? A mere syllogism indeed: tractable, uncomplicated, and obvious. However, social life is none of these things because people are in fact quite capable of crafting behaviors that are imperfect indexes of their inner beliefs. When people make polite conversation with a contemptible colleague, mask their desire for an attractive acquaintance, or feign interest in a friend's dreary problems, they are striking a thoughtful balance between the expression and inhibition of their true feelings. The self-regulation of overt behavior is the kingpin on which virtually all social interactions turn.

Yet, the same self-regulatory mechanisms that enable people to maintain peaceable relations with dreary friends and contemptible colleagues may also have their costs. One such cost lies in the way that self-regulation can affect the social perceiver's ability to draw accurate inferences about others. To understand how this happens it is useful to think of social inference not as a process but as a consortium of conceptually distinct processes (Gilbert, Pelham, & Krull, 1988 ; Quattrone, 1982 ; Trope, 1986). First, perceivers *categorize* or identify the actions of others ("Roger is criticizing Henry"), next they *characterize* or draw dispositional inferences about the actor ("Roger is a critical person"), and finally, they use information about the situational constraints that may have conditioned the actor's performance to *correct* or adjust those characterizations ("But Roger has been under a lot of stress lately, so perhaps he isn't really such a critical person after all").

Recent evidence suggests that these three sequential operations may differ in the amount of cognitive resources they require. Both categorization and characterization seem to be overlearned processes that generally occur outside of awareness and require relatively little conscious attention (Gilbert & Krull, 1988 ; Johnson, Jemmott, & Pettigrew, 1984 ; Kassin & Baron, 1985 ; Lowe & Kassin, 1980 ; McArthur & Baron, 1983 ; Newtson, 1980 ; Winter & Uleman, 1984 ; Winter, Uleman, & Cunniff, 1985 ; cf. Bassili & Smith, 1986 , and Higgins & Bargh, 1987 , pp. 376—378). Correction, on the other hand, seems to be an effortful, deliberate form of reasoning that requires a significant expenditure of cognitive resources (Gilbert, in press ; Gilbert et al., 1988). It is easy to see that Roger is overly critical but much harder to consider the many reasons why he may have behaved critically despite the lack of a critical predisposition.

If characterization is easy and correction is hard, then perceivers who become involved in activities that usurp their cognitive resources should have little difficulty characterizing others but should show a diminished ability to correct those characterizations. Gilbert et al. (1988) explored this hypothesis by showing subjects several video clips of an anxious-looking woman who was said to be discussing a series of anxiety-provoking topics (e.g., her sexual fantasies) with a stranger. Whereas control subjects simply watched these video clips and made ratings of the woman's dispositional anxiety, a second group of subjects saw these same clips while simultaneously rehearsing some word strings (a resource-consuming task). The results revealed that this second group of cognitively busy subjects were more likely than control subjects to conclude that the woman in the video clips was dispositionally anxious; in other words, they were less likely to correct their characterizations by using information about the anxiety-provoking topics that the woman was discussing. Interestingly, this was true despite the fact that the word strings that they had rehearsed were the anxiety-provoking topics themselves. Cognitively busy subjects, then, failed to correct their characterizations despite the fact that they were rehearsing the very information necessary for such a correction. Gilbert et al. concluded that resource-consuming activities such as cognitive rehearsal can impair the correction phase of the social inference process.

It seems likely that the self-regulation of overt behavior is, like the rehearsal of word strings, a resource-consuming enterprise. In general, self-regulation requires that a person

envision the potential consequences of different courses of action and then inhibit those actions whose imagined effects are undesirable. Even if the person decides to perform no action at all (e.g., when one refrains from correcting a spouse's inaccurate anecdote in the presence of guests) both the decision and the subsequent inhibition may require considerable cognitive work (see Gray, 1975 ; Logan & Cowan, 1984 ; Pennebaker & Chew, 1985 ; Waid & Orne, 1981 ; Wegner, Schneider, Carter, & White, 1987). It follows that perceivers who devote substantial resources to the regulation of their own behavior should be unable to devote these same resources to the difficult task of understanding those with whom they are interacting. As such, self-regulating perceivers may be able to categorize the behavior they see and characterize the target in terms of that behavior, but they may be unable to use situational constraint information to correct their characterizations. Experiment 1 sought to test this hypothesis.

Experiment 1

Method Overview

Subjects watched seven silent clips from a videotape of an interview. In five of the seven clips, the interviewee (the target) appeared quite depressed and unhappy. Half the subjects learned that in these five clips the target had been answering sadness-inducing questions, whereas the remaining subjects learned that in all seven clips the target had been answering happiness-inducing questions. Half of the subjects in each of these conditions were required to perform a self-regulation task (i.e., to avoid looking at words that appeared at the bottom of the video screen) while viewing the tape, and the remaining subjects were not. After viewing the tape, subjects rated the target's dispositional sadness, attempted to recall the questions that the target had answered, and attempted to recall and recognize the words that had appeared at the bottom of the screen.

Subjects

A total of 51 female students at the University of Texas participated to fulfill a requirement in their introductory psychology course. Only native speakers of English were eligible to participate.

Instructions

On arrival at the laboratory, subjects were greeted by a male experimenter who gave them a brief oral introduction to the experiment, provided them with complete written instructions, and then escorted each subject to a separate cubicle (equipped with video monitor and video camera) where the subject remained for the duration of the experiment.

The written instructions explained that, as part of a project on "the role of nonverbal behavior in the interviewing process," subjects would watch seven short clips from a videotape of an interview that had ostensibly taken place earlier in the year. Subjects were told that several female students had earlier been invited to participate in a

"personal history interview" and that those who participated had been asked to answer seven questions. The subjects were told that they would be seeing a short (approximately 15 s) clip from each of the seven responses provided by one randomly chosen interviewee (the target). It was explained that during the interview the camera had been placed behind the interviewer, and thus only the target would be visible on the screen.

Situational Constraint Information

Subjects were told that to protect the privacy of the target, the videotape would be shown without any sound. However, subjects were also told that they would be able to tell which of the seven questions the target was answering in any given clip because the question that the target was answering would appear on the screen for 10 s in written titles just prior to each clip.

Half of the subjects were randomly assigned to the *sad questions* condition. In this condition five of the seven titles indicated that the target was answering sadness-inducing questions (e.g., "Describe a time when your parents made you feel unloved"). In each of these five instances, the target hung her head, averted her gaze, and appeared generally listless and depressed. In the two remaining instances, the titles indicated that the target was answering happiness-inducing questions (e.g., "What is the nicest thing your parents ever did for you?"), and in these instances the target smiled, maintained eye contact with the interviewer, and appeared relatively chipper. The remaining subjects were assigned to the *happy questions* condition. In this condition subjects saw precisely the same behavioral episodes as were seen by subjects in the sad questions condition. However, in this condition all seven of the titles indicated that the target was answering happiness-inducing questions.

Thus, in the sad questions condition the target's apparent sadness could logically be attributed to the nature of the questions she was answering (i.e., the situational constraints on her behavior). However, in the happy questions condition, the identical behavior could not logically have been caused by the nature of the questions. Thus, behavior in the happy questions condition warranted a more dispositional explanation than did the identical behavior in the sad questions condition. The questions and the target's behavior in each of these conditions are shown in Table 1 .

Parafoveal Optiscope

In each subject's booth was a camera that appeared to be connected to an electronic device. Subjects were told that this device "is a parafoveal optiscope which will be recording your eye movements as you watch the film. Later on, we will be able to use this information to see just what parts of the film you looked at." Ostensibly, this information would tell the experimenter what features of the target's nonverbal behavior most subjects found informative. All subjects watched the film with their heads held at a constant distance from the screen and floor by a chin rest, ostensibly so that the optiscope would be able to track their eye movements. To enhance the credibility of this deception, all subjects went through a bogus calibration procedure in which they fixated on a point

of light as it moved across the screen while the experimenter purportedly adjusted the optiscope. The purpose of this deception is explained shortly.

To-Be-Ignored Words

All subjects were told that "in another condition of the experiment, other subjects will be asked to learn a list of words while they watch the film that you are about to see. As you'll notice when the film starts, a bunch of everyday words will appear and disappear on the screen." During the film, 38 common one-syllable nouns (e.g., tree, chair, sky) appeared, one at a time, at the bottom of the screen. After appearing, each word moved slowly upward and then disappeared before arriving at the middle of the screen. These words appeared during presentation of the target's behavior, but not during presentation of the written titles. The true purpose of these words (hereinafter referred to as the to-be-ignored, or TBI, words) is explained shortly.

Self-Regulation Manipulation

All subjects were told that their primary task was to form an impression of the target with particular emphasis on how dispositionally happy or sad she was. Subjects were told "we want you to tell us what kind of *person* the participant is, not just how she is *acting* ." Subjects were reminded that "most people seem unhappy from time to time, but only some of them are actually depressed sorts of people. Others are actually happy, optimistic types in their day-to-day lives, but may be temporarily unhappy because of a certain situation they are in."

Half of the subjects were randomly assigned to the *unregulated* condition. Subjects in this condition were told that as the film proceeded they should not concern themselves with the TBI words "because they are not relevant to the condition of the experiment that you are in. You are not in the word-learning condition and thus you can ignore these words." The remaining subjects were assigned to the *self-regulated* condition. These subjects were given the above instructions with regard to the TBI words, but in addition were told that the optiscope would not work if the subject moved her eyes too much. Subjects were told, "keep your eyes focused on the woman's (the target's) face during the film clips and do not under any circumstances look down at the words that are appearing and disappearing at the bottom of the screen. If you do accidentally look at one of the words, look away as quickly as possible so that the optiscope can readjust its alignment."

Thus, subjects in the unregulated condition were told that they *could* ignore the TBI words, whereas subjects in the self-regulated condition were told that they *should* ignore the TBI words. We assumed that these latter subjects would perform a conscious act of behavioral self-regulation that would usurp cognitive resources, thus leaving them unable to use the situational constraint information (i.e., the happiness- or sadness-inducing questions) when making judgments about the target.

Dependent Measures Perceived dispositional sadness.

Before the experiment began, subjects were allowed to familiarize themselves with the dispositional sadness measures. These measures required subjects to rate the target on four 13-point bipolar scales that were anchored at the endpoints with the phrases (a) is generally a happy (unhappy) sort of person, (b) is probably pretty cheerful (somewhat depressed) much of the time, (c) is generally a light-hearted (troubled) sort of person, and (d) probably has an optimistic (a pessimistic) outlook on life.

Recall of interview questions.

After seeing the videotape, subjects completed the dispositional sadness measures just described. Next, subjects were asked to recall each of the seven interview questions that the target had ostensibly answered.

Recall of TBI words.

Next, subjects were asked to recall as many of the TBI words as they possibly could. We assumed that self-regulated subjects might feel reluctant to admit that they had seen the TBI words because they had originally been forbidden to do so. Thus, all subjects were told, "It is a well-known scientific fact that some people are capable of what we call *parafoveal processing* . That is, even words that are not precisely focused on the fovea of the eye can nonetheless be read, understood, and recalled." We hoped that this claim would attenuate any reluctance on the part of self-regulated subjects to admit that they remembered the TBI words. In addition, we assumed that the parafoveal optiscope would function like a bogus pipeline (Jones & Sigall, 1971) and would lead all subjects to believe that the experimenter already knew whether or not they had looked at the TBI words.

Recognition of TBI words.

Next, subjects were given an alphabetized list of 72 common one-syllable nouns, 38 of which were the TBI words. Subjects were asked to circle those words that had appeared in the film and were encouraged to guess if they were unsure. Finally, all subjects were carefully probed for suspicion and fully debriefed.

Results and Discussion Perceived Dispositional Sadness

The data from two subjects who failed to follow instructions were discarded prior to analysis. Subjects' ratings of the target on the dispositional sadness scales were averaged to create a dispositional sadness index. The internal reliability of this index (coefficient alpha = .91) was increased by deleting one of the items (namely, lighthearted/troubled). A 2 (question type: happy or sad) × 2 (subject's behavior: regulated or unregulated) analysis of variance (ANOVA) revealed only the predicted Question Type × Subject's Behavior interaction, F 1, 45 = 3.96, p = .053 . As Table 2 shows, unregulated subjects considered the target more dispositionally sad when she appeared sad while answering happiness-inducing rather than sadness-inducing questions, F 1, 45 = 4.91, $p < .05$. In other words, unregulated subjects used the situational constraint information to correct (i.e., to

discount or to augment) their characterizations of the target. In contrast, self-regulated subjects considered the target to be equally sad, regardless of the type of questions she had ostensibly answered ($F < 1$). Self-regulated subjects, then, showed no evidence of inferential correction (i.e., of either discounting or augmenting).

Recall of Questions

The results are consistent with our hypothesis that self-regulation can usurp resources and thereby impair the ability to use situational constraint information. However, is it possible that self-regulated subjects did not use the situational constraint information simply because they did not have the information in memory? A question-recall index was computed by awarding subjects two points if they remembered the question verbatim, one point if they remembered the gist of the question but not its precise wording, and zero points if they remembered the question incorrectly or not at all. Thus, subjects could score from 0 to 14 points on the question-recall index.

A 2 × 2 ANOVA performed on the question-recall index revealed only an unimportant main effect of question type, F 1, 45 = 6.05, $p < .02$, such that subjects in the sad questions condition remembered more questions ($M = 11.04$) than did subjects in the happy questions condition ($M = 9.83$). Thus, although self-regulated subjects were less likely than unregulated subjects to use the situational constraint information, they were equally likely to have that information in memory, F 1, 45 = 2.03, $p > .16$.

Recall and Recognition of TBI Words

The number of TBI words that subjects recalled was submitted to a 2 × 2 ANOVA , which revealed no significant effects (all F s < 1.3). A similar ANOVA performed on d' (the signal-detection index of recognition accuracy) also revealed no significant effects (all F s < 1.9). Self-regulated and unregulated subjects were equally accurate in their recognition and recall of the TBI words.

It is worth noting that both correct recall and recognition of the TBI words were quite low: On average, subjects recalled a mere 4% of the TBI words ($M = 1.6$ words) and recognized the TBI words at about the chance level (mean $d' = 0.4$). Apparently, both self-regulated and unregulated subjects successfully ignored the TBI words, and it is somewhat ironic in this regard that self-regulated subjects evidently made a conscious (and costly) effort to avoid performing a behavior that, had they been unregulated, they would easily have avoided anyway.

Experiment 2

The results of Experiment 1 suggest that the active self-regulation of behavior can impair some ongoing cognitive operations (correction) without impairing others (characterization). Subjects who were merely asked to regulate their gaze drew more dispositional inferences about a target whose behavior was situationally induced than did subjects who were allowed to gaze where they wished. This operationalization of self-

regulation mirrors a variety of real world scenarios: For example, when people interact with handicapped individuals, pregnant women, provocatively clad members of the opposite sex, or others of interesting or unusual appearance, they may find themselves inhibiting a natural desire to stare (cf. Langer, Taylor, Fiske, & Chanowitz, 1976). In such cases, individuals may spend so much effort regulating their gaze that they are unable to perform the resource-limited operations that accurate social inference requires. It is worth noting that ours may have been a somewhat conservative test of our hypothesis in that subjects were asked to inhibit behaviors that they surely had little desire to perform in the first place. The inhibition of behaviors that beg for expression is probably a more demanding task.

Nonetheless, a far more common sort of behavioral self-regulation pervades our daily interactions. The hallmark of human relations is that a person's behavior does not always faithfully express his or her true feelings, beliefs, and desires. One may find a coworker unusually attractive and yet, for ethical reasons, be determined not to let it show; or one may find an in-law particularly distasteful but display the warmth and interest that familial politics require. It is likely that displaying false feelings toward an interaction partner can require a good deal of cognitive work. First, one must profess the behaviors associated with the attitude one is pretending to hold, and second, one must inhibit the natural tendency to behave in ways associated with one's true attitude. Although people may occasionally be compelled to display a false disdain, it is certainly more common for people to feign admiration for those whom they inwardly despise. Few of us can afford the luxury of candor toward those myriad gatekeepers whose souls we detest but whose grace we require.

In Experiment 2 we attempted to model this common social predicament. We asked female subjects to use nonverbal behavior to ingratiate a male confederate whose political opinions the subject was eliciting in an interview. The confederate was always constrained to give conservative responses to the interviewer's questions. We predicted that those subjects who ingratiated a dislikable (rather than a likable) confederate would devote a great deal of conscious attention to masking their true feelings, and would thus be particularly unlikely to use situational constraint information when drawing inferences about the confederate's true political opinions. In contrast, we predicted that observers (who were not required to ingratiate the confederate) would be just as likely to use situational constraint information when drawing inferences about the opinions of a dislikable as of a likable confederate.

Method Overview

Pairs of female subjects met a male confederate who behaved in either a likable or dislikable manner. One of the subjects (the interviewer) interviewed the confederate (the responder) about his opinions on several political issues while the other subject (the observer) watched the interview. Both subjects were told that regardless of his true opinions, the responder would read experimenter-generated conservative responses to the political interview questions. In addition to asking questions, the interviewer was

instructed to ingratiate the responder during the interview. Finally, all subjects attempted to estimate the responder's true political attitudes.

Subjects

A total of 46 female students at the University of Texas participated to fulfill a requirement in their introductory psychology course. Only native speakers of English were eligible to participate.

Likability Manipulation

Subjects were invited to participate in an experiment on "how people encode and decode communications." Two female subjects arrived at the laboratory and were escorted by a male experimenter to a waiting area. The experimenter explained that a third subject (actually a male confederate) was due to arrive momentarily. When the confederate arrived, the experimenter left the waiting area (ostensibly to check some equipment) and the confederate either engaged the 2 subjects in friendly conversation (the likable confederate condition) or ignored them thoroughly (the dislikable confederate condition).

The experimenter returned after 1 min and explained that before the experiment could begin he would have to retrieve some materials from a nearby building. In the dislikable condition the confederate was visibly annoyed with the delay (e.g., "You must be kidding. You mean we have to *wait* ?"), whereas in the likable condition the confederate assured the experimenter that the delay was "no sweat." After the experimenter left the waiting area for the second time, the dislikable confederate stormed out of the waiting area after telling the subjects, "If I'm not back on time, tell that jerk to start without me." Conversely, the likable confederate remained in the waiting area, behaved pleasantly, and offered both subjects a stick of gum during the experimenter's absence.

Interview Task

After returning with the materials that had ostensibly been retrieved from a nearby building, the experimenter escorted the subjects and the confederate to individual cubicles where each remained for the duration of the experiment.

One subject was randomly assigned the role of interviewer. From her cubicle the interviewer could see (and ostensibly could be seen by) the confederate via a closed circuit television system. The interviewer was told that her task would be to read seven questions to the confederate, who had been assigned the role of responder. It was explained that the responder would answer each of these questions by reading an experimenter-generated response, and that the interviewer's primary task was to diagnose the responder's true attitude on each of the issues. The experimenter claimed that this would help him learn "how people decode the communications of others." The interviewer was told, "Your task is not an easy one, because the responder will have no choice with regard to the response he gives. Rather, he will simply be reading answers that we have prepared."

Ingratiation Task

In addition, the interviewer was told that part of the experiment concerned "how people encode, or send, communications via nonverbal behavior." Thus, the interviewer was asked to use nonverbal means to ingratiate the responder during the interview. The interviewer was told, "You should do all those things that communicate liking, such as making eye contact, smiling, nodding, and so forth. Although *you* know that we have asked you to make yourself liked, we have not told the responder that you will be trying to make him like you."

The second subject was assigned the role of observer. From her cubicle the observer could see the interviewer and the responder on two adjacent video monitors. The observer was given the same information as was the interviewer. The observer was instructed to watch the ensuing interview with the goal of diagnosing the responder's true attitude on each of the seven political issues. The observer was, of course, informed that the interviewer would be trying to ingratiate the responder, and the interviewer was informed that the observer would be watching the interview.

Dependent Measures

Both subjects were given a sheet of paper containing seven 13-point bipolar scales anchored at the endpoints with the phrases *responder is opposed to* and *responder is in favor of* each of seven political issues (namely, military spending, fighting communism, legalized abortion, nuclear weapons, gun control, capital punishment, and school prayer). After the interviewer asked the first question and after both subjects heard the responder's first response, the subjects completed the first scale (i.e., estimated the responder's true attitude on the first issue). Subjects continued in this way until the last response had been given and the last scale was completed.

After the interview was concluded, subjects completed several additional 13-point bipolar scales that measured (a) how much the subject liked the responder and (b) how much the subject thought the responder liked the interviewer. In addition, observers completed a scale that measured (c) how much they thought the interviewer liked the responder, and interviewers completed a scale that measured (d) how difficult they had found the ingratiation task to be. In addition, during the interview the experimenter observed and rated (e) the interviewer's behavior on a scale that ranged from *not likable* to *very likable* . At the end of the experiment both subjects were carefully probed for suspicion and fully debriefed.

Results and Discussion Manipulation Checks

Subjects were asked how much they liked the responder. A 2 (role: interviewer or observer) × 2 (responder's behavior: likable or dislikable) ANOVA performed on these ratings revealed a main effect of responder's behavior, F 1, 42 = 30.87, $p < .0001$. As expected, subjects liked the likable responder a great deal more than they liked the dislikable responder (M s = 9.0 and 5.4, respectively).

In addition, interviewers were asked how difficult the ingratiation task had been. As expected, those who ingratiated a dislikable responder found the task much more difficult than did those who ingratiated a likable responder (*M* s = 7.6 and 2.5, respectively), *t* 21 = 4.36, *p* < .001 . Given the differential difficulty of the two tasks, it is not surprising that the experimenter's ratings of the interviewer's behavior revealed that interviewers behaved in a somewhat more likable way toward the likable than the dislikable responder (*M* s = 6.9 and 5.3 respectively), *t* 21 = 1.67, *p* < .10 , one-tailed. Interviewers who ingratiated a dislikable confederate apparently put forth exceptional effort with unexceptional results.

Ratings of Responder's Political Attitudes

We assumed (and interviewers agreed) that ingratiating a dislikable responder would be more cognitively demanding than ingratiating a likable responder. Thus, we predicted that interviewers would draw more dispositional inferences about a dislikable than about a likable responder, whereas observers would draw similar inferences about a likable and a dislikable responder.

A cursory examination of subjects' ratings of the responder's political attitudes on each of the seven issues revealed a substantial range of correlations between items. Thus, rather than averaging the items to form a perceived conservatism index, the items were submitted to a principal-components factor analysis whose solution was rotated by the varimax method. As Table 3 shows, three factors with eigenvalues greater than 1 explained 68% of the variance in subjects' ratings. For each subject, a factor score on each of these three factors was computed (using BMDP4M) and these scores were submitted to a 2 (role: interviewer or observer) × 2 (responder's behavior: likable or dislikable) ANOVA .

Factor 1: Political conservatism.

The first factor explained the greatest amount of variance and appeared to describe the political conservatism dimension that we had hoped to capture. This factor was defined by conservative attitudes toward nuclear weapons, fighting communism, capital punishment, and military spending, a constellation of issues that seems to revolve around threats to political and personal liberty and prescribed action toward threatening agents (e.g., executing criminals, fighting leftists, etc.).

A 2 × 2 ANOVA performed on this political conservatism index revealed only the predicted interaction, *F* 1, 42 = 5.86, *p* < .02 . As Table 4 shows, observers attributed similar degrees of conservatism to both the likable and dislikable responder, *F* 1, 42 = 1.89, *ns* , but interviewers attributed more conservatism to the dislikable than to the likable responder, *F* 1, 42 = 7.45, *p* < .01 . These findings are consistent with our hypothesis that being friendly to someone whom we actually dislike is a difficult task that impairs the ability to use situational constraint information and thereby facilitates dispositional attributions about that person's behavior.

One curious aspect of the results, however, is that the observers' ratings appear to be about midway between the ratings of the two groups of interviewers rather than being equivalent to the ratings of those who interviewed a likable confederate. One possible reason for this is that observers were instructed to watch two different people on two different video monitors, and thus had to split their attention between the interviewer and the responder. It is likely that observers found themselves evaluating the interviewer's ingratiation attempts as well as evaluating the responder's political opinions. As such, observers may have, in fact, been more cognitively busy than those subjects who interviewed a likable confederate.

Regardless of the reason for this marginal elevation of the observers' ratings, the important point is that no differences emerged between observers' ratings of the likable and of the dislikable responder, F 1, 42 = 1.89, *ns* . This is quite important because it suggests (a) that disliking a person does not in and of itself lead to dispositional attributions about that person's behavior, and (b) that there is no simple tendency to consider dislikable persons politically conservative. Rather, the expression of false feelings and the inhibition of true feelings appear to be critical ingredients in the recipe for uncorrected characterization.

Factor 2: Traditional morality.

The second factor appeared to describe a traditional morality dimension and was defined by conservative attitudes toward school prayer and legalized abortion. Conservative beliefs on these issues reflect adherence to traditional moral positions (e.g., belief in God, sanctity of life). A 2 × 2 ANOVA performed on this traditional morality index revealed only a marginal main effect of responder's behavior, F 1, 42 = 3.07, $p < .09$. The likable responder was seen as holding more traditional moral values (M = 32.4) than was the dislikable responder (M = 27.4).

This finding is of special interest. Recall that before they heard the responder speak, subjects already had some information about him, namely, that he was a rather nice or particularly nasty young man. Thus, subjects may have predicted the responder's attitudes on moral issues from their knowledge of the responder's likability alone. They may have attributed high morality to the likable responder and low morality to the dislikable responder regardless of what the responder said. Indeed, it seems reasonable to suspect that our female subjects would believe that a man who is ill-mannered toward young ladies in a waiting room is likely also to be contemptuous of unborn children and God, despite his claims to the contrary.

Correlational analyses provide strong evidence for this interpretation. If interviewers estimated the responder's morality strictly on the basis of the responder's likability, then we would expect a positive correlation between these ratings. On the other hand, if their estimates of the responder's political conservatism were influenced strictly by the amount of cognitive work they performed, then we would expect a positive correlation between these ratings. Of course, simple correlations cannot address this issue because likability and cognitive work were purposefully confounded in our experiment: Those interviewers

who ingratiated a dislikable responder were also those who performed the greatest amount of cognitive work, r 21 = –.57, $p < .05$. Thus, partial correlations (beta weights) were computed to examine the independent effects of perceived likability and of cognitive work on interviewers' estimates of the responder's attitudes.

As the path diagram in Figure 1 shows, interviewers' ratings of the responder's likability were significantly correlated with their ratings of the responder's traditional morality (β = .47, $p < .05$), but not with their ratings of the responder's political conservatism (β = .12, *ns*). On the other hand, interviewers' ratings of how much cognitive work they performed (i.e., their answers to the question "How difficult was the ingratiation task?") were significantly correlated with their ratings of the responder's political conservatism (β = .55, $p < .05$) but not with their ratings of the responder's traditional morality (β = –.24, *ns*). It is important to note that these partial correlations reveal the independent effects of cognitive work and perceived likability on interviewers' attributions. As such, they argue that cognitive work increased the interviewers' tendency to infer conservatism (but not morality) from the responder's behavior, whereas perceived likability increased the interviewers' tendency to infer morality (but not conservatism) from the responder's behavior.

These findings make a second valuable point. As predicted, we found that observers attributed the same degree of conservatism to the likable and dislikable responders. One might argue that this null effect was a trivial result of the fact that observers' feelings toward the responder could have waned during the course of the experiment (i.e., that the likability manipulation did not affect observers as strongly as it affected interviewers). This argument seems implausible on two counts. First, at the end of the experiment, observers and interviewers reported liking the responder equally well. Second, observers' attributions of morality indicate that observers did distinguish between the likable and dislikable responders on the trait dimension that was conceptually linked to the responder's likability (i.e., the moral dimension). In other words, observers' judgments of morality were just as strongly affected by the responder's likable or dislikable behavior as were the interviewers' judgments of morality. Thus, the fact that observers saw no difference between the likable and dislikable responders in terms of political conservatism is a meaningful null effect.

Factor 3: Gun control.

The third factor was described by a single issue: gun control. Our debriefings made us realize that subjects knew little about this issue; in Texas, gun ownership (like other God-given rights) is not a matter of great controversy. In addition, most subjects were unable to tell which position (pro- or anti-gun-control) was the conservative one. As a result, ratings of the responder's attitude on this issue were essentially uncorrelated with ratings on other issues (mean r = .13) and thus composed a separate factor. Not surprisingly, a 2 × 2 ANOVA on these ratings revealed no significant effects (all F s < 1.2).

General Discussion

When we recently asked a group of colleagues what they would do if they found New York City suddenly depopulated, most were quick to describe a behavior forbidden them in daily life. One expressed a desire to dance naked on the infield at Yankee Stadium, another to drop melons from the top of the Empire State Building, and a third simply to stroll unmolested through Times Square. For each of our colleagues the imagined absence of others signaled a welcome reprieve from the seemingly endless variety of social strictures that normally constrain their every word and deed. For most people, it seems, social life is an extended exercise in the self-regulation of action.

The present studies speak for themselves and we will not belabor their results. Suffice it to say that in both experiments self-regulation seemed to impair subjects' abilities to use situational constraint information when interpreting another's actions. As such, self-regulators drew dispositional inferences about behaviors that could easily have been explained with reference to situational forces. This effect obtained when the to-be-regulated behavior involved glancing at an irrelevant word on a video monitor or when it involved expressing one's true disdain for another. Together, these studies suggest that self-regulatory efforts can affect the inferences perceivers draw about those with whom they interact.

It is important to recognize, however, that not all self-regulatory attempts will necessarily impair the social inference process in this way. Some social rules (e.g., three of the Ten Commandments) cross virtually all situational boundaries, and it seems likely that such ubiquitous restrictions are themselves overlearned and that their enforcement therefore requires little cognitive work. Other restrictions, however, do demand a high sense of occasion: The frank revelation of one's sexual proclivities or personal failures may be appropriate among intimates but not among strangers, and the taciturn civility that the college president expects may send precisely the wrong message to a trusted friend. In short, some self-regulations are more easily imposed than others, and the present studies suggest that social inference is most likely to be affected when perceivers attempt to oblige unfamiliar, temporary, or local norms.

Self-Regulation in Everyday Life Judging Powerful Persons

The present thesis may shed new light on some old problems. For example, people tend to draw dispositional inferences about those with superior status (Thibaut & Riecken, 1955), and about those on whom their outcomes depend (Berscheid, Graziano, Monson, & Dermer, 1979 ; Miller, Norman, & Wright, 1978 ; cf. Erber & Fiske, 1983). The first of these phenomena is usually explained with reference to the extraordinary diagnosticity of high status persons' behaviors ("Why would the boss compliment me if she didn't mean it?") and the second is explained with reference to our own need to predict and control those who may most profoundly affect our lives ("I wonder what he'd say if I asked for a raise?"). The present studies suggest a somewhat simpler explanation of both effects: Powerful people may simply impair one's ability to process information.

When one interacts with powerful individuals one may take special pains to say and do the right things and to avoid the wrong ones, either because one's fortune depends on it or

merely because of the embarrassment that powerful persons may so easily inflict. As a result, perceivers may embrace dispositional explanations of powerful persons' behaviors simply because the perceiver's own self-regulatory actions have impaired his or her ability to draw accurate social inferences. This is not to say that standard motivational and attributional factors play no role in these outcome-dependence effects; indeed, a new professor may feel a pressing need to predict the dean and may also realize that a dean's actions toward a new professor are especially diagnostic in that they are relatively unencumbered by social norms. We wish merely to suggest that the basic mechanism outlined here may conspire with these other factors to achieve the same ends. Whenever situations place a premium on strategic self-presentation we may expect self-presenters to take others more or less at face value.

Judging the Out-Group

When individuals interact with members of other races, genders, and nationalities, they often draw more dispositional inferences from those behaviors than from the identical behaviors of their cohort. This phenomenon is generally explained with reference to a lack of cognitive complexity: People think of outgroups in simple ways and are thus overly influenced by the unrepresentative actions of out-group members (Linville & Jones, 1980 ; Quattrone & Jones, 1980). In addition, unfamiliarity with the norms that govern an out-group member's behavior may leave one unable to estimate the consensus that such actions enjoy. One may assume, for example, that a Japanese acquaintance is particularly reserved without realizing that such reserve is demanded by Japanese custom and, as such, tells one little about the acquaintance as a unique individual. Again, our findings suggest that a different mechanism may also provide a good account of this effect.

Just as one often knows little about the norms that govern an out-group member's behavior, one is often unsure of the norms that should govern one's own behavior in the company of out-group members. If one does not choose to err on the side of reticence then one must at least stay on guard against potentially embarrassing parapraxes. In either case, interaction with out-group members may cause individuals to be overly cognizant of their own actions, which may in turn impair their ability to draw accurate inferences about the out-group members with whom they are interacting. Once again, this explanation is not at odds with traditional explanations; rather, it complements them.

Judging the Self

Although the present studies deal only with inferences about others, they encourage speculation about the role of cognitive busyness in inferences about the self. Decades of research on cognitive dissonance and self-perception have shown that people often come to believe that which they once merely professed; that is, people often do not use information about the situational constraints on their own behavior (see Bem, 1972 , and Wicklund & Brehm, 1976 , for reviews). Although this phenomenon has begotten many elaborate theoretical explanations, our studies suggest a very simple reason for the effect.

Everyone is, from time to time, induced to mask some sentiments and to express others. This demanding enterprise may cause one to fail to consider the very forces that led one to lie in the first place. Indeed, the more elaborate and difficult the lie, the more likely one should be to believe that one believes it. Thus, a person who is subtly induced to tell another that a mundane peg-turning task is in fact quite thrilling may at one level recognize that this lie was coerced. Nonetheless, because telling such a lie with apparent conviction requires a great deal of effort and attention, the person may not be able to use the coercion information when drawing inferences about his or her own beliefs (cf. Festinger & Carlsmith, 1959). In short, responding to social pressures may leave us too busy to think about them.

Liars and Other Self-Regulators

Our cultural wisdom tells us that liars are particularly suspicious sorts who project their own deceptive aims onto others; as such, we might expect liars to be relatively immune to others' lies. Nonetheless, our studies suggest that this will not always be the case. As every philanderer knows, generating a cogent account of one's doings is a great deal more demanding than merely describing an event that has actually transpired (DePaulo, Stone, & Lassiter, 1985), and keeping the real and proffered accounts separated in memory may itself be an effortful chore (Johnson & Raye, 1981). Lying, then, can be a strenuous labor that leaves liars especially vulnerable to others' lies. One can envision the ironic consequences of this proposition: When two colleagues praise each other despite the enmity they feel, or when two unfaithful lovers explain their whereabouts the night before, each may be a successful perpetrator only at the cost of being an unwitting pawn. The more attention one's own actions require, the less attention one can spend drawing inferences about the actions of others.

If this is so, then why does our culture teach us that liars are less rather than more susceptible to lies? Although the cognitive busyness hypothesis predicts greater susceptibility, there are several other factors that may attenuate or even reverse this effect. First, a successful lie must often be told again and again: As some presidential hopefuls have found, a person who tells one reporter that he was born in Iowa and another that he was born in Massachusetts is rarely invited to live in Washington. By telling a particular lie on many occasions the lie may become well practiced, thus easing the cognitive burden of generating fresh lies. Second, liars may not trust themselves. That is, liars may fail to correct their characterizations of others, but they may also place little faith in these characterizations simply because they have learned from experience how easily one individual's view of another can be manipulated (cf. Toris & DePaulo, 1985). As a result, habitual liars may to some extent disregard the inferences they draw during interaction.

Finally, when people tell lies they often leak their true feelings through nonverbal behavior. As Gilbert and Krull (1988) have shown, when verbal and nonverbal behaviors are at odds, cognitive busyness can uniquely impair the perceiver's ability to draw dispositional inferences from verbal behavior. Thus, cognitively busy perceivers (of which liars are one sort) often rely on the nonverbal behavior of others, and in some cases

this may allow them to be more accurate lie detectors. On balance, then, all of these points suggest that liars and other self-regulators may well achieve a limited immunity to the lies of others.

Reprise

"My thinking," wrote William James (1890 , p. 333), "is first and last and always for the sake of my doing." True enough. Although modern psychology has emphasized the extent to which thinking guides doing, it has paid significantly less attention to the converse relation: How does our doing affect how and what we think? There is some irony in the answer offered here. Because social behavior often has such profound and inexorable consequences, people generally take great care to say and do the right thing at the right time. Yet, to the extent that such self-regulatory endeavors require conscious attention, they may predictably impair the individual's ability to understand the very persons for whom they were done in the first place.

References

Bassili, J. N. & Smith, M. C. (1986). On the spontaneity of trait attribution: Converging evidence for the role of cognitive strategy. *Journal of Personality and Social Psychology, 50*, 239-245.

Bem, D. J. (1972). Self-perception theory.(In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 6, pp. 1—61). New York: Academic Press.)

Berscheid, E., Graziano, W., Monson, T. & Dermer, M. (1979). Outcome dependency: Attention, attribution, and attraction. *Journal of Personality and Social Psychology, 34*, 978-989.

DePaulo, B. M., Stone, J. L. & Lassiter, G. D. (1985). Deceiving and detecting deceit.(In B. R. Schlenker (Ed.), *The self in social life* (pp. 323—370). New York: McGraw-Hill.)

Erber, R. & Fiske, S. T. (1983). Outcome dependency and attention to inconsistent information. *Journal of Personality and Social Psychology, 47*, 709-726.

Festinger, L. & Carlsmith, J. (1959). Cognitive consequences of forced compliance. *Journal of Abnormal and Social Psychology, 58*, 203-210.

Gilbert, D. T. ((in press). In J. S. Uleman & J. A. Bargh (Eds.),). Thinking lightly about others: Automatic components of the social inference process.

Gilbert, D. T. & Krull, D. S. (1988). Seeing less and knowing more: The benefits of perceptual ignorance. *Journal of Personality and Social Psychology, 54*, 193-202.

Gilbert, D. T., Pelham, B. W. & Krull, D. S. (1988). On cognitive busyness: When person perceivers meet persons perceived. *Journal of Personality and Social Psychology, 54*, 733-740.

Gray, J. A. (1975). *Elements of a two-process theory of learning.* (New York: Academic Press)

Higgins, E. T. & Bargh, J. A. (1987). Social cognition and social perception. *Annual Review of Psychology, 38*, 369-425.

James, W. (1890). *The principles of psychology* ((Vol. 2). New York: Holt)

Johnson, J. T., Jemmott, J. B. & Pettigrew, T. F. (1984). Causal attribution and

dispositional inference: Evidence of inconsistent judgments. *Journal of Experimental Social Psychology, 20*, 567-585.
Johnson, M. K. & Raye, C. L. (1981). Reality monitoring. *Psychological Review, 88*, 67-85.
Jones, E. E. & Sigall, H. (1971). The bogus pipeline: A new paradigm for measuring affect and attitude. *Psychological Bulletin, 76*, 349-364.
Kassin, S. M. & Baron, R. M. (1985). Basic determinants of attribution and social perception.(In J. Harvey & G. Weary (Eds.), *Attribution: Basic issues and applications* (pp. 37—64). New York: Academic Press.)
Langer, E. J., Taylor, S. E., Fiske, S. T. & Chanowitz, B. (1976). Stigma, staring, and discomfort: A novel stimulus hypothesis. *Journal of Experimental Social Psychology, 12*, 451-463.
Linville, P. W. & Jones, E. E. (1980). Polarized appraisals of out-group members. *Journal of Personality and Social Psychology, 38*, 689-703.
Logan, G. D. & Cowan, W. B. (1984). On the ability to inhibit thought and action: A theory of act control. *Psychological Review, 91*, 295-327.
Lowe, C. A. & Kassin, S. M. (1980). A perceptual view of attribution: Theoretical and methodological implications. *Personality and Social Psychology Bulletin, 6*, 532-542.
McArthur, L. Z. & Baron, R. M. (1983). Toward an ecological theory of social perception. *Psychological Review, 90*, 215-238.
Mencken, H. L. (1924). *Prejudices: Fourth series.* (New York: Knopf)
Miller, D. T., Norman, S. A. & Wright, E. (1978). Distortion in person perception as a consequence of the need for effective control. *Journal of Personality and Social Psychology, 36*, 598-607.
Newtson, D. (1980). An interactionist perspective on social knowing. *Personality and Social Psychology Bulletin, 6*, 520-531.
Pennebaker, J. W. & Chew, C. H. (1985). Behavioral inhibition and electrodermal activity during deception. *Journal of Personality and Social Psychology, 49*, 1427-1433.
Quattrone, G. A. (1982). Overattribution and unit formation: When behavior engulfs the person. *Journal of Personality and Social Psychology, 42*, 593-607.
Quattrone, G. A. & Jones, E. E. (1980). The perception of variability within in-groups and out-groups: Implications for the law of small numbers. *Journal of Personality and Social Psychology, 38*, 141-152.
Thibaut, J. W. & Riecken, H. R. (1955). Some determinants and consequences of the perception of social causality. *Journal of Personality, 24*, 113-133.
Toris, C. & DePaulo, B. M. (1985). Effects of actual deception and suspiciousness of deception on interpersonal perceptions. *Journal of Personality and Social Psychology, 47*, 1063-1073.
Trope, Y. (1986). Identification and inferential processes in dispositional attribution. *Psychological Review, 93*, 239-257.
Waid, W. M. & Orne, M. T. (1981). Cognitive, social, and personality processes in the physiological detection of deception.(In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 14, pp. 61—106). New York: Academic Press.)
Wegner, D. M., Schneider, D. J., Carter, S. R. & White, T. L. (1987). Paradoxical effects of thought suppression. *Journal of Personality and Social Psychology, 53*, 5-13.

Wicklund, R. A. & Brehm, J. W. (1976). *Perspectives on cognitive dissonance.* (Hillsdale, NJ: Erlbaum)
Winter, L. & Uleman, J. S. (1984). When are social judgments made? Evidence for the spontaneousness of trait inferences. *Journal of Personality and Social Psychology, 47,* 237-252.
Winter, L., Uleman, J. S. & Cunniff, C. (1985). How automatic are social judgments? *Journal of Personality and Social Psychology, 49,* 904-917.

This research was supported by National Science Foundation Grant BNS-8605443 to Daniel T. Gilbert.
We thank John Bargh, Bella DePaulo, Jamie Pennebaker, Janet Spence, Bill Swann, Dan Wegner, and several anonymous reviewers for their comments on an earlier version of this article. We also thank Susie Palmer for serving as the interviewee in Experiment 1 and Terry Grover for his painstaking portrayal of the responder in Experiment 2.
Correspondence may be addressed to Daniel T. Gilbert, Department of Psychology, University of Texas, Austin, Texas, 78712.
Received: August 24, 1987
Revised: February 29, 1988
Accepted: March 8, 1988

Table 1
Interviewer's Questions and Target's Responses by Condition

Sad questions condition	Happy questions condition	Target's response
Describe a time when your parents made you feel unloved.	What is your fondest childhood memory?	Sad
What was the biggest disappointment you ever had?	What was the most fun you ever had celebrating a birthday?	Sad
Has a close friend ever deserted you when you needed them?	What is the nicest thing your parents ever did for you?	Sad
Who is the funniest person you know?	Who is the funniest person you know?	Happy
Have you ever had someone close to you die?	What do you like most about your best friend?	Sad
What has been your proudest moment in school?	What has been your proudest moment in school?	Happy
What was the most painful time of your childhood?	What future goals or plans do you most look forward to?	Sad

Table 2
Subjects' Ratings of Target's Dispositional Sadness

	Subject's behavior	
Question type	Self-regulated	Unregulated
Happy	8.64 (12)	9.95 (13)
Sad	9.09 (11)	8.51 (12)
Difference	−0.45	1.44*

Note. Higher values indicate greater perceived dispositional sadness on a 1 to 13 scale. Cell *n*s appear in parentheses.
* $p < .05$.

Table 3
Loadings of Seven Political Issues on Three Factors

Issue	Political conservatism (Factor 1)	Traditional morality (Factor 2)	Attitude toward gun control (Factor 3)
Nuclear weapons	.844	—	—
Fighting communism	.750	—	—
Capital punishment	.628	—	—
Military spending	.635	—	—
School prayer	—	.766	—
Legalized abortion	—	.713	—
Gun control	—	—	.931
% variance explained	30	23	15

Note. Factor loadings < .5 have been omitted for presentation.

Table 4
Subjects' Ratings of Responder's Political Conservatism

Responder's behavior	Subject's role	
	Interviewer	Observer
Dislikable	35.6 (11)	28.6 (11)
Likable	24.8 (12)	31.3 (12)
Difference	10.8*	−2.7

Note. Values are factor scores transformed for presentation ([score × 10] + 30). Higher values indicate greater perceived political conservatism (i.e., greater perceived correspondence between responder's statements and dispositions) on a 0 to 60 scale. Cell *n*s appear in parentheses.
* $p < .05$.

Figure 1. The relations between interviewers' estimates of (a) the difficulty of the ingratiation task (cognitive work), (b) the responder's likability, (c) the responder's political conservatism (Factor 1), and (d) the responder's traditional morality (Factor 2). (The work/conservatism and work/morality correlations are independent of likability; the likability/conservatism and likability/morality correlations are independent of work. All values are beta weights, except the conservatism/morality and work/likability values, which are zero-order correlations. Values shown in boldface are $p < .05$ [$df = 21$]; all other values are $p > .25$ [$df = 21$].)