

Overhelping

Daniel T. Gilbert

Department of Psychology University of Texas at Austin

David H. Silvera

Department of Psychology University of Tromsø

ABSTRACT

Overhelping occurs when one attempts to spoil an observer's impression of a performer by explicitly helping the performer achieve a goal, thereby inviting the observer to attribute the performer's success to the help. The results of 4 experiments suggest (a) that people are most likely to overhelp when they believe that their interventions will be ineffective but will be considered effective by observers and (b) that when either of these beliefs is wrong, the strategy will backfire. The results point to an *intervention principle* that predicts how and when people may most effectively influence a performance so as to shape observers' inferences about the performer.

This research was supported by a grant from the National Institute of Mental Health (NIMH; 1-R01-MH49613-01), a Research Scientist Development Award from the NIMH (1-KO2-MH00939-01), and a predoctoral fellowship from the National Science Foundation Research Fellowship Program (GER-9253937). The support of these institutions is gratefully acknowledged. We thank Roy Baumeister, Bella DePaulo, James Tedeschi, Abraham Tesser, and James Shepperd for their comments on this article, and Julianne Baldwin, Rachel Glina, Jay Hook, and Holly Lewis for their help in the execution of these experiments. An earlier version of this article benefited greatly from the comments of the late Ned Jones.

Correspondence may be addressed to Daniel T. Gilbert, Department of Psychology, University of Texas, Austin, Texas, 78712.

Electronic mail may be sent to gilbert@psy.utexas.edu

Received: September 14, 1994

Revised: July 30, 1995

Accepted: October 13, 1995

There are different ways of assassinating a man—by pistol, sword, poison, or moral assassination. They are the same in their results except that the last is more cruel.—Napoleon I, *Maxims* (1804—1815)

It is good to win. After all, successful performances usually bring tangible rewards, such as wealth, power, and a rash of new friends. But if it is good to win, it is even better to be seen as a winner—a person whose success is due to superior ability and exemplary effort rather than birthright, assistance, or dumb luck. Lottery winners and ketchup heirs may be enviably affluent, but they are rarely invited to write for the *Wall Street Journal* or join a president's circle of economic advisors. Sweet success is sweeter, then, when it is seen as a measure of the person who attains it. But how can we tell when a performer's triumph should be taken as a sign of insight and cunning and when it should not? Attribution theories suggest that all performances are due to some combination of temporary situational and enduring dispositional factors and that those who wish to estimate the magnitude of the latter must use the discounting principle to "subtract out" the contribution of the former ([Heider, 1958](#) ; [Jones & Davis, 1965](#) ; [Kelley, 1967](#) ; cf. [Gilbert & Malone, 1995](#) ; [Higgins & Winter, 1993](#)). The investor's recent killing on the market may have been due to her uncanny instinct and management bravado (which led her to buy shares just before the stock split), or it may have been due to family connections and old school ties (which furnished her with inside information about an impending merger). If her financial coup is to be taken as a measure of her business acumen, an observer must first remove the effects of facilitating agents such as Uncle Standish and the "old girl network" down at the club. The more one comes to suspect that her successful performance was facilitated by others, the less confident one will be about her investment skills (see also [Jones, 1989](#)).

The Strategy of Overhelping

For the most part, the Uncle Standishes of the world probably mean well. The teachers, parents, lovers, coaches, mentors, and colleagues who facilitate our performances are often trying to increase the probability that we will reap the tangible benefits of success. Nonetheless, as an attributional analysis suggests, such facilitation may come at a cost. Insofar as observers (i.e., third parties, helpers, or even performers themselves) attribute a performer's success to the intervention of a helper, the performer may be deprived of credit for his own successful performance. In the hands of the attributional sophisticate, help can therefore be a powerful weapon—a subtle means of spoiling a performer's reputation for excellence while appearing to mean her well. The mother who insists on helping her son tie his shoes even though he is perfectly capable of doing so himself may undermine the child's conception of his capabilities. If the mother finds the dependence of her child particularly gratifying and fears that it will be diminished by the sense of self-efficacy that successful shoe-tying produces, she may (consciously or unconsciously) render unnecessary assistance. This habit will not only prolong the mother's gratification by forestalling the child's recognition of his own abilities, but it will also probably strike an impartial observer as nurturing, considerate, and kind. It is easy to imagine how the provision of assistance, or *overhelping*, might be used by mothers, husbands, athletes, businesspeople, and even well-placed uncles, to rob a performer of the fruits of her own achievements.

Overhelping is a way of spoiling a person's reputed ability with acts of apparent charity. But do people really use help to inflict harm? The evidence is mixed. Some research

suggests that when people wish to tarnish another's image to protect their own sense of self-worth, they tend to use an *underhelping* strategy that involves the provision of inadequate assistance rather than an *overhelping* strategy that involves the provision of superfluous assistance. For example, [Tesser and Smith \(1980\)](#) found that when male participants and their best friends competed on an important, ability-linked task, participants provided their friends with relatively worthless clues. That is, they attempted to spoil their friend's reputation for competence by inhibiting—and *not* by facilitating—his performance. Indeed, when people want to make others appear foolish or inept, it seems only natural that they will look for ways to ensure the other's failure. If a dancer wants his competitor to come off like a klutz during her audition, it seems more natural for him to trip her than for him to help her execute a graceful pirouette and then hope that the casting director discounts the flawless maneuver. People probably think about pushing their enemies toward disaster before they think of pulling them toward victory.

Nonetheless, some research suggests that people do occasionally stumble on the overhelping strategy. [Shepperd and Arkin \(1991\)](#) had participants perform an ability-linked task and allowed participants to determine which of several varieties of music a confederate would hear while he or she performed the same task. Some of the pieces of music were ostensibly performance-enhancing and others were ostensibly performance-inhibiting. When participants were led to doubt their ability to outscore the confederate and were led to believe that the experimenter would be directly comparing their score with the confederate's score, they were especially likely to expose the confederate to performance-enhancing—and *not* performance-inhibiting—music. One interpretation of these results is that participants attempted to spoil the confederate's reputed ability while masquerading as kindly disc jockeys. Under some circumstances, then, people do seem willing to facilitate, rather than inhibit, the performances of those whose apparent competence they wish to spoil.

Determinants of Overhelping

Both underhelping and overhelping are strategies for undermining a performer in the eyes of an observer, and there is some empirical evidence for each. What then determines which of these strategies a person will elect to use on a particular occasion? Surely there are a host of sociological, motivational, cognitive, and personality variables that should encourage different people to use different strategies at different times. But we suspect that the most fundamental determinant of one's choice of strategies is the relation between the behavioral and attributional consequences of help—that is, the relation between what help *does* to a performance and what observers *think* help does to a performance. We argue that overhelping is usually a risky strategy, that people tend to use that strategy only when the typically high risks are atypically low, and that people calculate these risks by gauging the behavioral and attributional consequences of their helping interventions. But before we do, we should offer a caveat. Our focus is on the interpersonal tactics that people may use to tarnish an observer's impressions of a target. As such, our discussion presumes that people are occasionally motivated toward such dishonorable ends. No one can say whether the desire to undermine the apparent competence of another is rare or common, but it seems reasonable to assume that as a

motive it appeals to some of the people some of the time. In any event, our theorizing is agnostic in this regard, and our aim is simply to describe and explore a strategy by which people may make others appear incompetent should they so desire. Whether, when, and how often people so desire is a question whose answer is beyond the scope of the present work.

The Intervention Principle

When a person helps another, she intervenes in the other's performance, and all such interventions have two kinds of consequences. First, when an intervention alters the probability that a performer will succeed it is said to have behavioral consequences. Help usually facilitates performances, and thus these consequences are usually *behavioral benefits* for the performer. Second, when an intervention alters the probability that the performance will be taken by observers as a clear indicator of the performer's ability it is said to have attributional consequences. Help usually leads observers to discount performances, and thus these consequences are usually *attributional costs* to the performer. We suspect that the relation between behavioral benefits and attributional costs primarily determines when people who are motivated to undermine a performer will overhelp or underhelp. Specifically, such people should overhelp only when they believe that attributional costs are greater than behavioral benefits—in other words, only when the possible improvement in the performance is more than offset by the denial of credit to the performer.

In the language of conditional probability, behavioral benefits may be defined as [p (success/help) – p (success/no help)], and attributional costs may be defined as [p (credit/success/no help) – p (credit/success/help)]. As such, we can state as a general principle that when a person wishes to undermine the perceived competence of a performer, then p (help) = f(attributional costs – behavioral benefits). This simple formula (which, for the sake of convenience, we refer to as the *intervention principle*) asserts that overhelping is likely to occur to the extent that behavioral benefits are smaller than attributional costs. So, for example, if the stockbroker has already crafted a good investment strategy from her own careful analysis of the market, then the malevolent Uncle Standish would be wise to make a public display of giving her precisely the information that she has already gathered. In such a case, his help will have few behavioral benefits (after all, his niece would have invested well without the advice), but it may have severe attributional costs (her fellow brokers may credit the accomplishment to her avuncular muse). On the other hand, if the stockbroker has not done her homework, then good advice may have behavioral benefits (it may save her a humiliating loss) that will not be offset by attributional costs. In such a case, an ill-intentioned uncle should remain silent and let his niece flounder—or perhaps directly undermine her by slipping her a note containing the very worst advice. This sort of intervention will probably have substantial behavioral costs (his niece will lose a bundle) and few attributional benefits (her fellow brokers will know nothing of the secret note, will not realize she was sabotaged, and will thus attribute her bad fortune to her lack of financial skill). In short, a wicked uncle's choice of strategies should be determined by the relation between the attributional and behavioral consequences that he expects his actions to have.

Risks of Overhelping

People should overhelp when they expect the behavioral benefits of help to be offset by the attributional costs. The problem is that people can rarely be certain about how costly or beneficial a particular intervention will be: Uncle Standish may believe that his niece already has excellent market information, and thus he may expect his sage advice to be superfluous. But that expectation may be dashed. One can imagine, for example, that receiving a recommendation from her uncle that validates her own analysis may increase the niece's confidence so much that she invests \$1 million rather than the \$1,000 that she had originally intended and thus reaps a windfall. Her fellow brokers may recognize her uncle's role in her accomplishment but may nonetheless be impressed by the steely nerve of someone who bets \$1 million. This example merely illustrates what the intervention principle codifies: There are two risks inherent in any overhelping attempt, namely, that the helper may underestimate the behavioral benefits of his assistance (which we call *the behavioral-prediction problem*) and that he will overestimate its attributional costs (which we call *the attribution-prediction problem*).

Each of these problems entails its own psychology, because the helper's prediction of behavioral benefits and attributional costs are determined by very different factors. Predictions of the behavioral benefits of help depend on the helper's beliefs about the *performer's* ability and motivation to succeed at the task without help (how smart, well-informed, and hardworking is the niece?) as well as the *actual* value of the help rendered (how useful is the financial advice?). If a performer's ability and motivation are either extremely low or extremely high, she may not profit from help, and thus the behavioral benefits of help will be small. In such instances, the risks of overhelping should be diminished. On the other hand, the helper's estimates of the attributional costs of help depend on the helper's beliefs about the *observer's* ability and motivation to make accurate attributions (how thoughtful are the niece's fellow brokers?), as well as the *apparent* value of the help (how useful do the other brokers consider the advice to be?). If the kind of help rendered is thought by most people to be particularly useful and the observers are particularly thoughtful, then such observers are likely to discount a successful performance, and the attributional costs of help will be increased. If the overhelper incorrectly estimates either the performer's drive and ability to succeed or the observer's drive and ability to use the discounting principle, overhelping may backfire.

Both the behavior-prediction problem and the attribution-prediction problem are serious and suggest that overhelping is a strategy fraught with peril. The overhelper must have a keen sense of both the actual and apparent value of the help he provides as well as a keen sense of both the performer's and observers' motivations and capacities. The wise overhelper should offer assistance that does not actually affect performance (either because it is objectively worthless or because the performer's success or failure are already assured by the performer's own ability and motivation) in the presence of an observer who believes the aid facilitated the performance and who can and will consider that fact when making an attribution. Alas, this is by no means an easy prescription to follow.

The Present Studies

The foregoing analysis suggests some circumstances under which overhelping will and will not be an effective strategy. On those occasions when ordinary people do wish to damage another's reputation for competence, can they use the complex logic of this analysis to guide their own interpersonal behavior? The present studies were designed to provide a preliminary answer to this question. In particular, in our first study we attempted to learn whether people understand the logic of overhelping well enough to provide assistance to a performer whom they wished to undermine. We created a set of circumstances that our analysis suggested would be conducive to overhelping, and then we watched to see whether our participants would invent and implement the strategy. In our second set of studies we attempted to answer three more specific questions about the tactic of overhelping: (a) does the ill-intentioned helper's estimate of behavioral benefits and attributional costs control his or her decision to use underhelping or overhelping strategies? (b) does overhelping actually spoil the performer's reputation for competence in the eyes of the observer? and (c) when does overhelping backfire?

Experiment 1

Method Overview

Participants watched videotapes of two men and one woman separately answering questions about their attitudes toward a variety of social issues. The videotapes were designed so that (a) one man was likable and one man was dislikable, and (b) the men and the woman had virtually identical attitudes on four topics. Participants were told that the two men and the woman would be returning to the laboratory later, that the woman would hear the men answer new questions about the same set of topics, and that the woman would then choose one of the men for a date. Participants were asked to choose the new questions that would be posed to each of the men in the woman's presence. Participants were given a choice between neutral questions that merely asked the man to report his true attitude and helpful questions that blatantly encouraged the man to respond as both he and the woman had responded before. We expected that participants would overhelp the dislikable man by asking him helpful questions, thus causing the woman to suspect that his answers were elicited by the question rather than provoked by his true beliefs.

Participants

Participants were 31 female undergraduates at the University of Texas at Austin who took part in the experiment to fulfill an obligation in their introductory psychology course.

Instructions

Participants reported to the laboratory, where they were greeted by a female experimenter who ushered them to a private cubicle where they remained for the duration of the

experiment. Participants were given both written and oral instructions that explained that the experimenter was studying "how the people who introduce dating partners to one another can affect the likelihood that a woman will accept or reject a man's invitation for a date." Participants were told that earlier in the semester, two male students and one female student had volunteered to take part in an experimental simulation of a television show called "The Dating Game." These 3 students (hereinafter referred to as *the men* and *the woman*) had ostensibly stated that they were not at present involved in a romantic relationship and that they wanted to meet new people. The men and the woman had ostensibly come to the laboratory at different times and had been interviewed, and each of these interviews had been videotaped with their permission. Participants were told that before the experimental task was fully explained, they would see a videotape of the men and the woman being interviewed. In fact, the men and the woman were confederates who gave scripted responses, and no simulation of "The Dating Game" ever took place.

Videotaped Interviews

Participants saw three videotaped interviews. Participants first saw the woman being interviewed. The videotape showed a male interviewer making a series of 12 statements and, after making each statement, asking the woman to verbally agree or disagree with the statement and to briefly explain her position. Next, participants saw one of the men being interviewed in the same manner. Finally, participants saw the other man being interviewed in the same manner. The interviewer made the same 12 statements in the same order in each of the three videotaped interviews. These 12 statements concerned mundane preferences ("Tofu is good to eat"), public policies ("Cigarette smoking should be banned in public places"), and personal beliefs ("There is nothing wrong with premarital sex"). It was made clear to participants that the three interviews had been conducted at different times, that none of the interviewees had ever met the others, and that none had seen or heard the other interviews.

Liking manipulation.

Responses to 4 of the 12 statements were designed to lead participants to dislike one man ("Mr. Awful") and to like the other ("Mr. Wonderful"). For example, when the interviewer stated "If women are working, then men should share in household tasks such as washing dishes and doing the laundry," the woman agreed and explained why. Mr. Awful answered the same question by saying "I disagree. My dad didn't do much work around the house and I'm pretty much hopeless. I can't iron a shirt without ironing in the wrinkles." Mr. Wonderful answered the question by saying "I agree. I don't think it's fair that women should do all the household tasks, especially if they're out there working too. I'd be willing to do my share." We assumed that our female participants would consider Mr. Awful's response less appealing than Mr. Wonderful's response. The remaining three statements concerned the appropriateness of the "obey clause" in the traditional marriage ceremony (the woman and Mr. Wonderful considered it archaic, whereas Mr. Awful thought it was fine), whether people who are truly in love should still be attracted to others (the woman and Mr. Wonderful thought they should not, whereas Mr. Awful thought that "a healthy appetite for the opposite sex" was not "a bad thing"), and whether

welfare should be eliminated (the woman and Mr. Wonderful were sympathetic to the plight of welfare recipients, whereas Mr. Awful believed welfare recipients were "just scamming the system"). The liking manipulation statements were made second, third, eighth, and twelfth, respectively.

Critical statements.

The first, sixth, ninth, and tenth statements were designated as critical statements. On each of the topics addressed by these statements, the woman and both men were in complete agreement with each other. For example, the interviewer said "Heavy metal music is the coolest kind of popular music," and the woman disagreed and explained why. Mr. Awful said "I disagree. It's just a bunch of noise. It's music for people with rings in their noses and stuff." Mr. Wonderful said "I disagree. There may be a few good songs out there, but I think there are a lot more bad heavy metal bands than there are good ones. On the whole, it's loud, it's boring—just like the people who listen to it." The remaining three critical statements concerned a ban on public smoking (all were against a ban), premarital sex (all considered premarital sex permissible), and having children (all said they ultimately wanted to have children).

Manipulation checks.

When the videotapes were finished, participants were asked to report their impressions of the men and the woman on separate 7-point scales anchored at the endpoints with the phrases *do not like at all* and *like extremely well*. Participants then predicted on similar scales how well the woman would like each of the men. Finally, participants were asked "Which of the two males do you think should get a date with the female?" We expected participants to like Mr. Wonderful more than Mr. Awful, to predict that the woman would feel similarly, and to report that Mr. Wonderful should get the date.

The Question Choice Task

Participants were given further written instructions that explained the primary experimental task in detail. Participants were told that the men and the woman would be returning to the laboratory next semester to take part in an experimental simulation of a television show called "The Dating Game." Ostensibly, the 2 men would each be asked four questions by an interviewer, and the woman would observe their responses. The woman would ostensibly choose one of the men, and the experimenter would then send them on an expense-paid date to a fine restaurant and a theatrical or musical performance. Participants were told that "in a kind of roundabout way, you are going to introduce the two males to the female, and you are going to do this by determining exactly which questions the experimenter will ask the two male students during the Phase 2 interviews, which will be seen live by the female student."

Participants were then shown a list of the four critical topics. Three questions were printed beneath each topic. The first question was a neutral question, for example, "What's your attitude toward heavy metal music? Do you like it or dislike it?" The second

question was a helpful question that blatantly encouraged the man to respond as the woman would want him to respond and as he had responded during the earlier interviews. The question did this by explicitly stating the woman's preferences, for example, "The female student said she did not like heavy metal music very much. How do you feel about it? Do you like it or dislike it?" The third question was also a helpful one that encouraged the man to respond as the woman would want him to respond and as he had responded during the earlier interviews. The question did this by presenting arguments in favor of a particular response, for example, "Heavy metal music is not a very sophisticated form, and it is most popular among high school kids and skinheads. How do you feel about it? Do you like it or dislike it?" Participants were asked to determine which of these three questions would be posed to Mr. Wonderful and which would be posed to Mr. Awful when each returned to the laboratory. Participants made this decision for all four topics. Participants were supplied with written transcripts of the videotaped interviews to help them "remember what the three people were like, what attitudes and preferences they had in common, and so on" and were encouraged to "use these transcripts however you like." After participants completed the question choice task, they were probed for suspicion, debriefed, thanked, and ¹

Results and Discussion Omission of Data

One participant reported knowing one of the male confederates, and her data were not analyzed, thus leaving 30 participants in the data set.

Manipulation Checks

Participants liked the woman quite well ($M = 5.37$), they liked Mr. Wonderful much more than they liked Mr. Awful (M s = 5.77 and 2.90, respectively), $F(1, 29) = 112.65$, $p < .0001$, and they predicted that the woman would share that preference (M s = 5.83 and 2.20, respectively), $F(1, 29) = 267.30$, $p < .0001$. Indeed, 100% of the participants reported that they thought Mr. Wonderful should get the date. Apparently, the liking manipulation was quite powerful.

Question Choice Task

Helpful questions were those that blatantly encouraged the male student to respond as the woman would presumably have wanted him to respond. We created a *helping index* such that participants received 1 point for each helpful question they chose, and we submitted their total score on the index (which could vary from 0 to 4) to a 1×2 (male student: Mr. Wonderful or Mr. Awful) repeated measures analysis of variance (ANOVA), which revealed the predicted effect of male student, $F(1, 29) = 16.42$, $p < .0001$. Participants chose to pose more helpful questions to Mr. Awful ($M = 2.73$) than they chose to pose to Mr. Wonderful ($M = 1.37$). The number of participants who chose helpful questions for each of the men for each topic is shown in [Table 1](#).

Experiment 2a

Participants played the role of yenta with aplomb. When they wanted to keep a likable woman from dating a man whom they knew (but whom the likable woman did not know) was self-centered, immature, and arrogant, they helped the man precisely when help was least needed and thereby stole the thunder that they could not prevent. Participants apparently realized that the man would score points with the woman by expressing attitudes that were similar to hers, and so they chose to elicit those expressions with questions that begged—and begged loudly—for the right answer. Participants did not merely follow a primal impulse to "help the one you like." Rather, they thoughtfully considered the fact that the dislikable man's success was virtually assured by his true attitudes, and they concluded that the behavioral benefits of help would be nil, the attributional costs could be substantial, and thus they could best thwart the dislikable man by helping him.

Our theoretical analysis suggests that people are especially likely to use an overhelping strategy when behavioral benefits are nil, as they were in Experiment 1. Of course, we did not ask participants to choose questions to which Mr. Awful's natural response would have repelled, rather than attracted, the woman, and thus we cannot be sure that it was participants' estimates of behavioral benefits and attributional costs that induced them to use the overhelping strategy. In addition, we presumed that participants chose helpful questions to damage Mr. Awful's reputation, and postexperimental interviews suggested that this was indeed the explicit intention of many participants. Nonetheless, it seemed important to develop a paradigm in which participants could be *directed* to damage a performer's reputation for competence. In Experiment 2a we used a paradigm that was not only amenable to the manipulation of behavioral benefits, but one in which the overhelping strategy could actually be executed and its consequences observed.

Method Overview

Participants were told that a participant from an earlier experiment (the applicant) had taken an aptitude test for a job as a library research assistant. Participants were told that the results of a new aptitude test would be shown to another participant (the employer) who would decide whether the applicant deserved the job. The test consisted of anagrams that contained various numbers of hint letters. Participants saw the test results of either a normal or a brilliant applicant and were asked to enhance or spoil the applicant's chances of getting the job by supplying many or few hints on a new test that would ostensibly be given to the applicant and whose results would ostensibly be shown to the employer.

Participants

Participants were 20 male and 36 female undergraduates at the University of Texas at Austin who took part in the experiment to fulfill an obligation in their introductory psychology course.

Procedure

Participants arrived at the laboratory in small groups. The experimenter explained that he was investigating how people who write letters of reference for job applicants can influence an employer's impression of the applicant. After receiving a brief oral introduction, participants were seated at individual desks and were given complete written instructions. These instructions explained that participants would be asked to play a role in a simulation of the hiring process. Participants were told that the applicant had taken part in a previous experiment and that during that session the applicant had completed an aptitude test that consisted of several anagrams. Participants were told that in a later experiment another participant would be asked to play the role of employer, would be shown the results of a new anagram test, and would be asked to determine the suitability of the applicant for the job of library research assistant.

Manipulation of Applicant's Ability

Participants were shown a set of twelve 8-letter anagrams that had ostensibly been attempted by the applicant. Participants could see that for each anagram, the first zero to five letters of the solution had been provided as a "hint" (e.g., the anagram UNIMAOTN was followed by the 3-letter hint MOU_ _ _ _). Six of the anagrams were classified as difficult (i.e., 2 anagrams had no hint letters, 2 had one hint letter, and 2 had two hint letters), and 6 of the anagrams were classified as easy (i.e., 2 anagrams had three hint letters, 2 had four hint letters, and 2 had five hint letters). Participants were shown a page that contained the 12 printed anagrams, printed solutions to each anagram, and a handwritten solution for some of the anagrams. Participants were led to believe that the handwritten solutions had been written by the applicant and that the printed solutions had only afterward been printed on the sheet by the experimenter.

Half the participants were randomly assigned to see the test results of an exceptionally capable applicant ("Mr. Brilliant"). These participants saw correct, handwritten solutions to all 12 anagrams. The remaining participants were assigned to see the results of a moderately capable applicant ("Mr. Normal"). These participants saw correct, handwritten solutions for the 6 easy anagrams and no solutions for the 6 difficult anagrams. Participants were told that the lack of a handwritten solution indicated that the applicant had been unable to solve the anagram.

Anagram Test-Construction Task

Next, participants were shown three new 8-letter anagrams and their printed solutions and were told that the applicant had never seen these particular anagrams. Participants were told that the applicant would be returning to the laboratory for a new aptitude test that would consist of only these three new anagrams. They were also told that the applicant's performance on all three anagrams would be shown to the employer, who would then decide whether the applicant would be hired. Participants were asked to provide from zero to five consecutive hint letters (always starting with the first letter) for each of the three new anagrams. Half the participants were asked to provide hint letters with the goal of creating a test that would increase the applicant's chances of getting the job (the helping-goal condition), and the remaining participants were asked to provide hint letters

with the goal of creating a test that would decrease the applicant's chances of getting the job (the hindering-goal condition). At the end of the session, participants were probed for suspicion, debriefed, thanked, and dismissed.

Results and Discussion

The number of hint letters that were provided with each anagram was taken as an index of that anagram's difficulty (fewer hint letters indicated greater difficulty). We submitted the average difficulty of the three anagrams that the participant constructed for the applicant to a 2 (applicant: Mr. Brilliant or Mr. Normal) \times 2 (participant's goal: help or hinder) ANOVA. The analysis revealed only the predicted Applicant \times Participant's Goal interaction, $F(1, 52) = 14.41, p < .001$. As [Table 2](#) shows, participants attempted to hinder Mr. Normal by presenting him with relatively difficult anagrams (underhelping) and to help Mr. Normal by presenting him with relatively easy anagrams, $F(1, 26) = 11.97, p < .01$. However, participants who attempted to manipulate the employer's impression of Mr. Brilliant did just the opposite. These participants attempted to hinder Mr. Brilliant by presenting him with relatively easy anagrams (overhelping) and to help Mr. Brilliant by presenting him with relatively difficult anagrams, $F(1, 26) = 6.35, p < .02$. Clearly, participants expected that Mr. Brilliant's apparent competence would be undermined if his excellent performance were known to have been facilitated by hint letters. In short, participants tried to thwart the applicant by causing him to fail, but when they could not, they tried to rob him of credit for what seemed to be an unavoidable success.

Experiment 2b

The results of Experiment 2a support our suggestion that a helper's beliefs about the behavioral benefits of assistance will determine his or her choice of underhelping or overhelping strategies. As the intervention principle suggests, when attributional costs were held constant, the helpers' beliefs about the behavioral benefits of their assistance influenced their choice of tactics. Specifically, participants who were motivated to spoil a performer's reputation for competence helped those performers whose outcomes would probably have been unaffected by that help but hindered those performers whose outcomes they could influence. In short, participants seemed to know when overhelping was most and least likely to work, and they implemented and eschewed that strategy appropriately.

Participants in both of the foregoing experiments made a calculated choice of strategies. But were those calculations correct? Experiment 2a demonstrated that participants will overhelp a job applicant whose success is already assured, but is there any reason to believe that an employer would actually be taken in by this gambit? As we mentioned earlier, there are two very different problems confronting the would-be overhelper: (a) The behavior-prediction problem (help may inadvertently confer behavioral benefits that outweigh the attributional costs), and (b) the attribution-prediction problem (the observer may fall prey to the correspondence bias, may not consider the role that the help played in facilitating the performance, and the attributional costs may inadvertently be low). In

other words, overhelping should work best when performers benefit little from help that observers consider a lot. Conversely, overhelping should backfire when performers do benefit from help, when observers ignore the help that was provided, or both. In Experiments 2b and 2c we examined these circumstances, respectively.

Experiment 2b was an attempt to use the intervention principle to predict when the overhelping strategy devised by participants in Experiment 2a would, in fact, spoil an observer's impression of a real (rather than a hypothetical) test-taker. In Phase 1, participants of different ability levels were overhelped and underhelped as they took a pair of anagram tests. In Phase 2, a new group of observer participants was shown the results of some of these tests. The intervention principle suggests that observer participants in Phase 2 would, in fact, form particularly negative impressions of some of the performer participants who were overhelped in Phase 1 but not of others. Specifically, we predicted that help would have little or no behavioral benefit for extremely competent or extremely incompetent performer participants, and as such, overhelping would successfully damage their reputed abilities in the eyes of observer participants. We predicted that, on the other hand, help would confer behavioral benefits on moderately competent performer participants, and that as such, overhelping would backfire and enhance their reputed abilities in the eyes of observer participants.

Method Overview

In Phase 1, an exceptionally capable, a moderately capable, and a barely capable test-taker were selected from a group of participants who took a series of anagram tests. Each test-taker was then given two tests. On one test, the test-taker was overhelped (i.e., he or she was given many hint letters), and on one test the test-taker was underhelped (i.e., he or she was given few hint letters). In Phase 2, a new group of participants was shown the results of the anagram tests and was asked to assess the intellectual ability of the test-takers.

Phase 1: The Test-Takers

Participants. Participants were 32 male and 28 female undergraduates at the University of Texas at Austin who took part in the experiment to fulfill an obligation in their introductory psychology course.

Anagram tests. Participants arrived at the laboratory in groups and were seated at individual desks. The experimenter explained that the study was designed to explore the use of anagram tests as measures of intellectual ability and that each participant would be given two anagram tests. Participants were first presented with a set of five 8-letter anagrams and were allowed 10 min to solve them. Each of these anagrams contained four hint letters and was thus classified as an easy anagram. Participants who took this test were said to have been overhelped. Next, participants were presented with a new set of five 8-letter anagrams and were again allowed 10 min to solve them. Each of these anagrams contained only one hint letter and was thus classified as a difficult anagram.

Participants who took this test were said to have been underhelped. At the end of the session, participants were debriefed, thanked, and dismissed.

Test-taker selection. The anagram tests were scored and revealed a wide range of abilities among participants. The median number of anagrams correctly solved was five of the five easy anagrams and three of the five difficult anagrams. The 3 highest scoring participants correctly solved all 10 anagrams, and the lowest scoring participant correctly solved 1 of the easy anagrams and 1 of the difficult anagrams. On the basis of these scores, 3 participants were selected to play the role of test-taker in Phase 2. The best test-taker ("Mr. Brilliant") was randomly selected from the 3 top-scoring participants, the average test-taker ("Mr. Normal") was randomly selected from the 5 participants who scored at the median on both tests, and the worst test-taker ("Mr. Plodding") was the lowest scoring participant. It is worth noting that we could have concocted these results more easily than we extracted them from real participants, but we chose to do the latter so that there would be no question about the realism of the performances. So, for example, Mr. Plodding's performance is, by definition, a valid and representative example of a poor performance by a college student, and it was not somehow artificially manufactured to substantiate our hypotheses in Phase 2.

Phase 2: The Observers

Participants. Participants were 29 male and 36 female undergraduates at the University of Texas at Austin who took part in the experiment to fulfill an obligation in their introductory psychology course.

Instructions. Participants arrived at the laboratory alone and were greeted by an experimenter who ushered them to a cubicle where they remained for the duration of the experiment. The experimenter explained that the study was part of ongoing research into the testing process and that this study was designed to determine how observers assess the ability of test-takers.

Independent manipulations. Each participant was presented with the results of both an overhelped and an underhelped performance. These performances were always the performances of the same test-taker from Phase 1, but participants were told that they were the performances of two different individuals. One third of the participants saw the overhelped and underhelped performances of Mr. Plodding, one third saw the overhelped and underhelped performances of Mr. Normal, and the remaining participants saw the overhelped and underhelped performances of Mr. Brilliant. The order in which participants saw the underhelped and overhelped performances was counterbalanced across conditions. Thus, the experiment comprised a 2 (helping strategy: underhelped or overhelped) \times 3 (test-taker: Mr. Brilliant, Mr. Normal, or Mr. Plodding) completely factorial design, with the first of these being a within-subjects variable.

Dependent measures. Along with the test-takers' results, participants were given an evaluation form for each test-taker. The evaluation forms asked participants to estimate the following for each of the two test-takers: (a) The test-taker's IQ (participants were

told that 100 was the mean IQ), (b) the test-taker's verbal SAT score (participants were told that verbal SAT scores range from 200 to 800), (c) the test-taker's overall SAT score (participants were told that overall SAT scores range from 400 to 1600), (d) the test-taker's high school grade point average (GPA; on a 0—4 scale), and (e) the test-taker's college GPA (on a 0—4 scale). Next, the evaluation forms presented participants with a series of 15-point Likert-type scales on which they were to make various ratings of the test-takers. Participants were asked to evaluate (f) the likelihood that the test-taker would correctly solve a new 8-letter anagram with two hint letters (the scale ranged from *not at all likely* to *extremely likely*), (g) their certainty about the previous judgment (the scale ranged from *not at all certain* to *extremely certain*), (h) the test-taker's verbal intelligence (the scale ranged from *low verbal intelligence* to *high verbal intelligence*), (i) the test-taker's general intelligence (the scale ranged from *not at all smart* to *extremely smart*), and (j) how diagnostic the test was (specifically, participants were asked "How fair of a test of the test-taker's ability do you believe this particular test was?" They answered on a scale that ranged from *not at all fair* to *extremely fair*). Participants then completed a variety of exploratory measures, were probed for suspicion, debriefed, thanked, and dismissed.

Results and Discussion Perceived Ability Index

Eight of the ten dependent measures assessed the observer's impression of some aspect of the test-taker's ability. Those measures were (a) IQ, (b) verbal SAT, (c) overall SAT, (d) high-school GPA, (e) college GPA, (f) the likelihood that the test-taker would solve a new anagram, (h) verbal intelligence, and (i) general intelligence. Scores on these items were standardized and then combined to create a perceived ability index that had excellent reliability ($\alpha = .94$). We submitted this index to a 2 (helping strategy: underhelped or overhelped) \times 3 (test-taker: Mr. Brilliant, Mr. Normal, or Mr. Plodding) ANOVA, which revealed a main effect of test-taker, $F(2, 58) = 41.17, p < .001$; and a main effect of helping strategy, $F(1, 58) = 18.00, p < .001$; both of which were qualified by the predicted Test-Taker \times Helping Strategy interaction, $F(2, 58) = 16.61, p < .001$.

As the first row of [Table 3](#) shows, observers formed more negative impressions of Mr. Brilliant and Mr. Plodding when the test-takers were overhelped than when they were underhelped. On the other hand, observers formed more negative impressions of Mr. Normal when he was underhelped rather than overhelped. As participants in Experiment 2a would surely have predicted, providing help to a performer for whom the help has no behavioral benefits can indeed spoil an observer's impression of that performer. This was just as true when the nature of the performance was assured by the performer's high ability (Mr. Brilliant) as when it was assured by his low ability (Mr. Plodding). The data also show that attempts to overhelp backfired when the strategy was used on Mr. Normal, whose performance was indeed improved (from three correct to five correct solutions) when he was given help. Clearly, when help does have behavioral benefits, providing it will enhance rather than spoil a person's reputed ability.

Certainty and Diagnosticity

We submitted participants' ratings of the certainty of their impressions of the test-takers to a 2×3 ANOVA (as above), which revealed a main effect of test-taker, $F(2, 61) = 4.47, p < .02$, which was qualified by a Test-Taker \times Helping Strategy interaction, $F(2, 61) = 19.87, p < .001$. As the second row of [Table 3](#) shows, participants were less certain about their impressions of an overhelped than an underhelped test-taker when that test-taker was brilliant but were more certain when that test-taker was plodding. This finding makes good sense: Success at a difficult task (the underhelped Mr. Brilliant), and failure at an easy task (the overhelped Mr. Plodding) are, in fact, especially diagnostic indicators of ability, and participants seem to have been aware of this. It is worth noting that the means in every cell are near or above the midpoint of the scale. Although it is difficult to interpret the absolute meaning of any scale rating, the magnitude of these ratings at least suggests that participants were generally confident about their judgments of the test-taker's ability.

We submitted participants' ratings of the diagnosticity of the test to a 2×3 ANOVA (as above), which revealed a main effect of test-taker, $F(2, 61) = 4.59, p < .02$, which was qualified by a Test-Taker \times Helping Strategy interaction, $F(2, 61) = 12.69, p < .001$. As the third row in [Table 3](#) shows, diagnosticity ratings showed the same pattern of results as did the certainty ratings. Participants considered the easy test to be more diagnostic of Mr. Plodding's ability than was the difficult test and the difficult test to be more diagnostic of Mr. Brilliant's ability than was the easy test. Again, there is a great deal of wisdom in these reports.

Experiment 2c

Together, Experiments 2a and 2b make a simple point: People who wish to undermine the reputed ability of a performer will overhelp the performer when they believe their help will not have behavioral benefits, and this strategy can be quite effective. On the other hand, there are circumstances under which it will predictably backfire, namely, when the help confers significant behavioral benefits (as it did for Mr. Normal).

There is, of course, a second case in which overhelping should backfire. As we mentioned earlier, if overhelping is to succeed, then observers must make considerable use of the discounting principle—that is, they must disregard the fact of the performer's glaring success and focus instead on the external agents that facilitated that success. It is nearly a truism in social psychology that observers tend to concentrate on performances per se and often fail to take external facilitation into account ([Gilbert & Malone, 1995](#); [Jones, 1979](#); [Ross, 1977](#)). If observers fall prey to this correspondence bias—that is, if they fail to consider fully the effects of the helper's assistance—then the overhelping strategy may backfire. An evil uncle will spoil nothing but his own plan if his niece's fellow brokers do not see his hand in her accomplishments. Research has shown that people are especially unlikely to use the discounting principle (i.e., they are especially prone to correspondence bias) when they have limited cognitive resources ([Gilbert, 1989](#)). This suggests that when observers are simultaneously engaged in several resource-consuming tasks, they may fail to take into account the overhelper's facilitation of the performer's behavior and may judge the performer strictly on the basis of his success.

Although a lack of cognitive resources is usually thought of as a disadvantage, in this case it may keep observers from becoming the victims of an overhelper's manipulations by forcing them to ignore a situation that was, in a sense, engineered to mislead them.

In Experiment 2c we tested this hypothesis. In Experiment 2c, observer participants were or were not made cognitively busy (i.e., given an extra task to perform) and were then asked to form impressions of Mr. Brilliant. As Experiment 2b showed, overhelping is an effective way to spoil the impression that Mr. Brilliant makes on an observer. As such, it seemed to us that Mr. Brilliant provided the best test case for the hypothesis that a normally effective overhelping strategy will backfire when observers are cognitively busy.

Method Overview

Participants were shown either the underhelped or the overhelped performance of Mr. Brilliant from Experiment 2b and were asked to evaluate his intellectual ability. In addition, some participants were asked to perform a tone-identification task while they reviewed the test results and evaluated Mr. Brilliant.

Participants

Participants were 59 male and 45 female undergraduates at the University of Texas at Austin who took part in the experiment to fulfill an obligation in their introductory psychology course.

Instructions

Participants arrived at the laboratory alone and were greeted by a male experimenter who ushered them to a cubicle where they remained for the duration of the experiment. The experimenter explained that the purpose of the experiment was to investigate the conditions under which people can perform more than one task at the same time. He told participants that mental functions are usually localized in particular cerebral hemispheres and that "creative processing" generally takes place in the right hemisphere, whereas "analytic processing" generally takes place in the left hemisphere. He explained that his hypothesis was that people could indeed perform two tasks concurrently as long as one of the tasks was "left brained" and the other was "right brained." This cover story was designed to eliminate possible demand characteristics by causing participants to embrace the opposite of the real hypothesis.

Participants then received written instructions. These instructions provided brief descriptions of two tasks that participants might be asked to perform: A "right-brained" tone-identification task and a "left-brained" person-evaluation task. After reading a description of each task, half the participants were informed that they would be asked to perform both the tone-identification and person-evaluation tasks at the same time (busy condition). The remaining participants were told that they would be asked to perform only the person-evaluation task (nonbusy condition).

Person-Evaluation Task

All participants were asked to review the results of a test taken by a student who had actually participated in a previous experiment. All participants were shown either the underhelped or the overhelped performance of Mr. Brilliant from Experiment 2b (i.e., five correct solutions of five 8-letter anagrams with one hint letter each and five correct solutions of five 8-letter anagrams with four hint letters each, respectively).

Dependent Measures

Along with the test results, participants were given an evaluation form that asked them to estimate the following items for the test-taker: (a) IQ, (b) verbal SAT score, (c) overall SAT score, (d) high school GPA, and (e) college GPA. Participants were also asked to estimate on a series of 15-point Likert-like scales (f) the likelihood that the test-taker could solve a new 8-letter anagram with two hint letters provided, (g) their certainty of the previous judgment, (h) the test-taker's verbal intelligence, (i) the test-taker's general intelligence, and (j) the diagnosticity of the test. Each of these was measured as in Experiment 2b. Finally, participants completed several exploratory measures.

Tone-Identification Task

Before busy participants were allowed to perform the person-evaluation task, the experimenter gave them training on the tone-identification task. Busy participants listened through headphones to an audiotape that played a sequence of three tones: A low-pitched tone, a medium-pitched tone (one octave higher than the low-pitched tone), and a high-pitched tone (two octaves higher than the low-pitched tone). Participants listened as the tones were twice played in this order (with durations and intertone intervals of approximately 1 s). The experimenter then asked participants if they felt they could distinguish the three tones from each other. All participants indicated they could.

The experimenter then told the busy participants that as they performed the person-evaluation task, they would also listen to an audiotape on which these three tones would be continuously played in a random order with random intertone intervals. Their job was to press the button on a hand-held counter whenever the high-pitched tone was followed by the medium-pitched tone and then the low-pitched tone. In other words, their job was to conduct an auditory search for the high-medium-low tone sequence. The experimenter then began the audiotape, and busy participants began performing the tone-identification task and the person-evaluation task simultaneously.

At the end of the experiment, all participants were probed for suspicion, debriefed, thanked, and dismissed.

Results and Discussion Perceived Ability Index

Eight of the ten dependent measures assessed the observer's impression of some aspect of the test-taker's ability. Those measures were (a) IQ, (b) verbal SAT, (c) overall SAT, (d)

high school GPA, (e) college GPA, (f) the likelihood that the test-taker would solve a new anagram, (h) verbal intelligence, and (i) general intelligence. Scores on these items were standardized and then combined to create a *perceived ability index* whose reliability was improved by the deletion of Items a and f ($\alpha = .78$). The deletion of these items did not change the pattern of results. We submitted this index to a 2 (helping strategy: underhelped or overhelped) \times 2 (observer's state: busy or nonbusy) ANOVA, which revealed only a main effect of helping strategy, $F(1, 99) = 14.49, p < .001$. Although the Helping Strategy \times Observer's State interaction was not reliable, $F(1, 99) = 2.33, p = .13$, the ANOVA interaction term probably provides an overly conservative test of our hypothesis, which is more appropriately tested by planned comparisons (Rosenthal & Rosnow, 1985; Rosnow & Rosenthal, 1995). Such comparisons support the hypothesis. Specifically, nonbusy participants did indeed form more negative impressions of Mr. Brilliant when he was overhelped than when he was underhelped, $F(1, 49) = 18.24, p < .001$, but busy participants did not, $F(1, 50) = 2.17, p = .15$ (see Table 4).

Certainty and Diagnosticity

We submitted participants' ratings of the certainty of their judgments to a 2×2 ANOVA (as above), which revealed no significant effects (all F 's < 1). The grand mean of these ratings is well above the midpoint of the scale and suggests that participants in all conditions felt quite confident about their judgments. Participants' ratings of the diagnosticity of the two tests were submitted to a 2×2 ANOVA (as above), which revealed only a main effect of helping strategy, $F(1, 100) = 4.22, p < .05$. Both busy and nonbusy participants considered the difficult test to be a more diagnostic measure of Mr. Brilliant's ability than was the easy test. Although the means in Table 4 suggest that the effect of helping strategy on perceived diagnosticity may have been greater in the nonbusy than in the busy condition, planned comparisons could not be used, because these differences were not theoretically predictable.

General Discussion

The external facilitation of a person's performance has both attributional and behavioral consequences, and these consequences may be independent. This simple fact has many intriguing ramifications, one of which is that interventions that enhance a person's performance may also undermine an observer's impression of the performer's ability. The foregoing studies demonstrate that people have an intuitive grasp of the probabilities that determine the effectiveness of overhelping—that is, they seem to know when help is likely to spoil a person's reputed ability and when it is likely to enhance it. This knowledge enables people who want to undermine another to do so, not by thwarting the other's performance, but by aiding and abetting it. Although such attempts are successful under some circumstances, they fail under others. Overhelping clearly backfires when its sanguine effects on a performer's behavior outweigh its insidious effects on the observer's attribution. As the intervention principle states, this is likely to be the case whenever behavioral benefits are particularly high (as they were for Mr. Normal in Study 2b), when attributional costs are particularly low (as they were for Mr. Brilliant in Study 2c), or both.

These studies merely scratch the surface of a potentially rich phenomenon that exists at the intersection of social cognition and interpersonal behavior and thereby draws on a large portion of social psychology's wisdom. It is not difficult to use the intervention principle to generate hypotheses about when people will overhelp, how the victims of overhelping might foil such efforts, and the circumstances under which observers will and will not be captured by the ploy. Nonetheless, before we discuss the implications of this work, it is important to reiterate a point that the work does *not* make. Like any set of laboratory investigations, this work does not tell us if or how often overhelping may occur in natural settings. It tells us that overhelping requires, at the very least, that a helper be motivated to make a performer appear incompetent and that a helper correctly predict that her intervention will have few behavioral benefits (which are primarily a function of the performer's native ability and the difficulty of the task) and considerable attributional costs (which are primarily a function of the observer's cognitive and motivational state). Although laboratory studies have shown that each of these effects *can* occur, no one knows how often they do occur and occur in tandem, and thus no one knows just how common overhelping itself is likely to be. It would be a serious misreading of our work to conclude that people are ruthless misanthropes bent on destroying the reputations of their competitors, that they typically do so in the guise of providing aid, or that the kindness of strangers is generally ill-intended. None of these implications can or should be drawn. Our point is simply that help can harm, that people recognize this, and that when they want to, they can press that recognition into service.

Just as Niccolò Machiavelli described his analysis of human motivation in the form of advice to a prince who might wish to subjugate his people, so too can we describe some of the implications of our work in the form of advice to a person who might wish to damage the apparent competence of another. In any given instance, help may be rendered in different ways and at different times, and the intervention principle provides counsel to the would-be overhelper about the optimal form and timing that his intervention should take.

How to Overhelp

All interventions can be construed in terms of their actual (behavioral) and apparent (attributional) effects. The less accurate an observer's understanding of such effects is, the more easily the overhelper should be able to use this inaccuracy to fool the observer. For example, overhelping should be particularly effective when it involves an intervention that actually degrades a person's performance but that is mistakenly thought by observers to enhance it. For example, observers may expect that a large parental support check will increase a college student's academic performance by allowing her more time for study, but in fact the check may decrease the student's intrinsic interest in school and thus prove an impediment to her success ([Lepper, Greene, & Nisbett, 1973](#)). Observers may assume that the threat of termination will facilitate an employee's performance by increasing his motivation, but in fact such a threat may actually hinder his performance by increasing his anxiety ([Baumeister, 1984](#)). In these and other cases, an intervention that appears to be facilitating is actually debilitating, and the overhelper who can

construct such interventions stands an improved chance of ruining the performer's reputed ability.

Such paradoxical interventions are surely the most effective weapons in the overhelper's arsenal but, because they require that observers be entirely wrong about how the world works, opportunities to provide them may be somewhat limited. On the other hand, there are surely many instances in which an observer's beliefs about the effects of an intervention on a performance are not wildly mistaken but are imperfectly calibrated. Students, for example, may believe that cramming for exams is a good way to boost their grades. Cramming does not debilitate performance inasmuch as it is surely better than no study at all, but such massed practice is clearly less effective than distributed practice ([Bray, 1948](#); [Gay, 1973](#)). If a misinformed student encourages his classmate to skip the weekly study sessions and instead concentrate his efforts on the night before the test, then similarly misinformed observers may expect more positive results than are actually warranted, and the exam-taker may be robbed of credit for whatever success he achieves ("Henry studied for 7 hours and only made a C+. What a dolt!").

The point here is simply this: Overhelping should be most effective when observers believe that the helper's intervention is more powerful than it really is, and the larger the discrepancy between the apparent and actual values of the intervention, the more effective overhelping should be. One potentially interesting direction for future research is to map people's naive beliefs about the power of common interventions onto the facts of the matter. It may turn out that people systematically underestimate the power of some interventions and overestimate the power of others, and the domains in which they do each of these may well suggest the circumstances under which overhelping is most likely to be observed.

When to Overhelp

Just as overhelpers must choose the particular form of help, so must they choose the moment at which to administer it. Helping interventions can occur at any one of several points, from the time a performer initially develops her skill to the time an observer learns about a performance (see [Aronson & Jones, 1992](#)). Three general points of intervention—before, during, and after a critical performance—suggest that there are three corresponding kinds of overhelpers, which we refer to as skill managers, task managers, and information managers.

A *skill manager* provides help by aiding a performer prior to the execution of a critical public performance, thus determining whether a performer's potential will be translated into a true ability to accomplish a variety of tasks. Coaches and teachers who train students prior to athletic competitions or standardized national tests are good examples of skill managers. A *task manager* provides help during the critical performance itself, essentially making the critical task easier to accomplish and less diagnostic (as did participants in our studies). Task managers determine whether a performer's ability will be translated into a successful performance. For example, a student who provides hints or clues during an exam is essentially altering the difficulty of the critical task and thereby

increasing her classmate's chances of success on that task alone. Finally, an *information manager* provides help after a critical performance by deciding which performances or portions of a performance will be relayed to an observer. Information managers determine whether a performance will be translated into a public observation. For example, a professor whose letter of recommendation documents a student's triumphs and carefully avoids any mention of his failures is helping the student by managing the information that the observer receives. Interventions at these three points entail very different considerations that may be powerful determinants of a person's decision to use or forego the overhelping strategy. Participants in our studies were task managers—that is, they influenced a performance by providing or withholding hints that made the task itself easier or more difficult to perform, and we have discussed the intricacies of such tactics at length. However, overhelping requires a very different set of tactics if it is to be performed effectively by either a skill manager or an information manager.

Considerations for Skill Managers

Effective overhelping requires that an observer witness a successful performance and that the observer attribute some portion of that success to the overhelper. Although skill managers can exert considerable influence on a variety of performances, there are at least two reasons why the skill manager's contributions to a particular performance are not likely to be recognized by observers (and therefore why the attributional costs of overhelping are likely to be low). First, skill managers typically provide help that is temporally removed from the critical performance itself; as such, their interventions may not be particularly salient for the observer. How often do cheering fans consider the role of the batting coach when the hometown slugger brings in the game-winning run? Second, skill managers do not influence a critical performance directly; rather, they influence a performance by influencing a performer's dispositions. As [Higgins and Winter \(1993\)](#) noted, attributional logic requires that observers discount the effects of *temporary* situational constraints (which affect the execution of behaviors) but not of *enduring* constraints (which promote the acquisition of dispositions). Even if the tennis champion who hails from an affluent family had the benefit of a lifetime of expensive lessons, the fact remains that she is an exceptional player, and thus the effects of the expensive lessons should not be "subtracted out" by an observer who wishes to assess her ability.

A skill manager's interventions, then, are relatively invisible and, under some circumstances, immaterial to the judgment of dispositions. As such, overhelping should be a particularly risky strategy for a skill manager to use. But the considerations that make overhelping a risky strategy are the same considerations that make underhelping a reliable one. Because the interventions of skill managers are generally invisible and occasionally immaterial, the skill manager who wishes to ruin a performer's apparent competence should probably provide hindrances rather than help. Even if the detrimental effects of these hindrances are well known, they are still unlikely to be considered by observers. For example, observers are unlikely to take into account the poor schooling and impoverished home environment of an inner-city child when they attempt to diagnose his ability to do math. Such socioeconomic and historical factors may have

exerted a dramatic impact on the child's performance, but they are far removed in time (observers may find it difficult to summon vivid images of barren ghetto landscapes when watching a child perform in a shiny new testing center) and, in some sense, immaterial to the judgment (whatever its cause, the fact of the child's quantitative weakness is indisputable).

Considerations for Information Managers

Information managers have a very different set of opportunities and concerns. The information manager is in the unique position of being unable to affect the critical performance itself; in other words, her interventions may have attributional costs, but they cannot have behavioral benefits. This fact can be played to the information manager's advantage. One of the risks of overhelping is that the helping intervention may inadvertently increase the performer's chances of success but, for any one of several reasons, observers will not be sufficiently cognizant of that fact. If the overhelper incorrectly calculates such probabilities, then his attempts at overhelping will backfire. The information manager is unique in not having to worry about this miscalculation problem. By the time the information manager comes on the scene, the performance is done, and all that remains is to transmit news about certain portions of that performance to an observer. An overhelping professor may, for example, write a glowing letter of recommendation in which she describes a student's mundane successes ("Sam received A's in all of his graduate courses") but fails to describe his more meaningful accomplishments ("Sam published three papers in prestigious journals and was the sole author on each"). Just as overhelping task managers may choose to make a task particularly easy when they suspect that the performer will succeed with or without their intervention, overhelping information managers may choose to *reveal* only information about performances on such easy tasks.

Information managers also differ from skill managers in that observers are quite unlikely to view their overhelping attempts in a positive light if discovered. In a sense, information managers operate on observers rather than on performers. If observers become aware that an information manager is manipulating their impressions of a performer, there is little room for a charitable interpretation of such action. A professor who is caught hiding his student's most outstanding achievements cannot, like the skill manager, claim that he was "only trying to help." In addition, because observers are especially eager to examine information that they believe is being kept from them ([Worchel & Arnold, 1973](#)), the overhelping information manager runs the risk of drawing attention to the very information she wishes to hide. For all of these reasons, information managers, like skill managers, are probably better off hindering than helping. The professor who includes information about a student's failures in her letter of recommendation can both sabotage the student's chances of success and claim that she was merely being fair in her assessment. Overhelping, then, can occur at any one of three stages, but our analysis suggests that it is most effectively used by task managers during the critical performance itself.

Coda

As Napoleon suggested, the flesh provides an adequate target for assassination, but it is the character that bleeds most profusely. If an individual has a reasonable understanding of attributional logic and can make a reasonable guess about the effects of her own interventions on another's performance, then she will have at her disposal a variety of strategies for spoiling the impression that a performer makes on an observer. One of these strategies—the provision of useless aid—is especially insidious in that it may undermine the performer's apparent competence while appearing to be driven by the most benevolent of motives. The foregoing studies show that people know when this strategy is likely to work and when it is likely to fail. Ordinary people are well-schooled in the art of killing with kindness, and it is incumbent on us to determine if, when, and how they put such skills to use in everyday life.

References

- Aronson, J. M. & Jones, E. E. (1992). Inferring abilities after influencing performance. (*Journal of Experimental Social Psychology*, 28, 277—299.)
- Baumeister, R. F. (1984). Choking under pressure: Self-consciousness and paradoxical effects of incentives on skillful performance. (*Journal of Personality and Social Psychology*, 46, 610—620.)
- Bray, C. W. (1948). *Psychology and military proficiency*. (Princeton, NJ: Princeton University Press)
- Gay, I. R. (1973). Temporal position of reviews and its effect on the retention of mathematical rules. (*Journal of Educational Psychology*, 64, 171—182.)
- Gilbert, D. T. (1989). Thinking lightly about others: Automatic components of the social inference process. (In J. S. Uleman & J. A. Bargh (Eds.), *Unintended thought* (pp. 189—211). New York: Guilford Press.)
- Gilbert, D. T. & Malone, P. S. (1995). The correspondence bias. (*Psychological Bulletin*, 117, 21—38.)
- Heider, F. (1958). *The psychology of interpersonal relations*. (New York: Wiley)
- Higgins, E. T. & Winter, L. (1993). The "acquisition principle": How beliefs about a behavior's prolonged circumstances influence correspondent inference. (*Personality and Social Psychology Bulletin*, 19, 605—619.)
- Jones, E. E. (1979). The rocky road from acts to dispositions. (*American Psychologist*, 34, 107—117.)
- Jones, E. E. (1989). The framing of competence. (*Personality and Social Psychology Bulletin*, 15, 477—492.)
- Jones, E. E. & Davis, K. E. (1965). From acts to dispositions: The attribution process in person perception. (In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 2, pp. 219—266). New York: Academic Press.)
- Kelley, H. H. (1967). Attribution theory in social psychology. (In D. Levine (Ed.), *Nebraska Symposium on Motivation* (Vol. 15, pp. 192—238). Lincoln: University of Nebraska Press.)
- Lepper, M. R., Greene, D. & Nisbett, R. E. (1973). Undermining children's intrinsic interest with extrinsic rewards: A test of the "overjustification" hypothesis. (*Journal of Personality and Social Psychology*, 28, 129—137.)

- Rosenthal, R. & Rosnow, R. L. (1985). *Content analyses: Forced comparisons in the analysis of variance*. (Cambridge, England: Cambridge University Press)
- Rosnow, R. L. & Rosenthal, R. (1995). Some things you learn aren't so: Cohen's paradox, Asch's paradigm, and the interpretation of interaction. (*Psychological Science*, 6, 3—9.)
- Ross, L. (1977). The intuitive psychologist and his shortcomings. (In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 10, pp. 173—220). New York: Academic Press.)
- Shepperd, J. A. & Arkin, R. M. (1991). Behavioral other-enhancement: Strategically obscuring the link between performance and evaluation. (*Journal of Personality and Social Psychology*, 60, 79—88.)
- Tesser, A. & Smith, J. (1980). Some effects of task relevance and friendship on helping: You don't always help the one you like. (*Journal of Experimental Social Psychology*, 16, 582—590.)
- Worchel, S. & Arnold, S. (1973). The effects of censorship and attractiveness of the censor on attitude change. (*Journal of Experimental Social Psychology*, 9, 365—377.)

1

We felt it would be unwise to ask participants to choose questions on topics with which Mr. Awful disagreed with the woman, or to include a hurtful question option. First, we feared that participants might feel that if they simply allowed Mr. Awful's disagreement on one topic to become evident to the woman, then no further sullyng of his reputation would be necessary. By allowing Mr. Awful to fail once, we may have eliminated the possibility that participants would resort to the overhelping strategy on other trials. Second, it is not entirely clear what the choice of a hurtful question would mean. For example, one might choose to ask a hurtful question ("Heavy metal music is just wonderful, don't you think?") in the belief that the man would answer as he truly felt ("No, I think it stinks") and thereby convince the woman that he was especially sincere in his statement, or in the belief that he would comply with the question and thereby convince the woman that he held a dissimilar attitude.

Table 1
Number of Participants Who Chose Helpful Questions for Mr. Wonderful and Mr. Awful in Experiment 1

Topic	Mr. Awful	Mr. Wonderful	<i>p</i> <
Heavy metal music	18	10	.07
Smoking ban	23	11	.003
Premarital sex	20	9	.006
Having children	21	11	.01

Note. All *p* values are from one-way repeated measure analyses of variance.

Table 2
Number of Hint Letters in Anagrams Presented to Applicants in Experiment 2a

Participant's goal	Applicant's ability	
	Normal	Brilliant
Help	2.65 (17)	1.69 (16)
Hinder	1.70 (11)	3.06 (12)

Note. Higher numbers indicate easier anagrams (i.e., more hint letters provided). Numbers in parentheses are cell *ns*.

Table 3
Observers' Impressions of Test-Takers in Experiment 2b

Measure	Test-taker's ability					
	Brilliant		Normal		Plodding	
	Underhelped	Overhelped	Underhelped	Overhelped	Underhelped	Overhelped
Perceived ability index	0.908 (21)	0.131 (21)	0.059 (20)	0.378 (20)	-0.488 (20)	-1.084 (20)
Certainty	12.95 (22)	8.59 (22)	9.67 (21)	9.48 (21)	7.29 (21)	9.95 (21)
Diagnosticity	8.09 (22)	5.09 (22)	7.33 (21)	7.76 (21)	4.14 (21)	6.10 (21)

Note. Numbers in parentheses are cell *ns*. Differences in *ns* between measures are due to failure of some participants to complete all measures.

Table 4
Observers' Impressions of the Brilliant Test-Taker in Experiment 2c

Measure	Observer's state			
	Nonbusy		Busy	
	Underhelped	Overhelped	Underhelped	Overhelped
Perceived ability index	0.293 (26)	-0.383 (26)	0.182 (25)	-0.108 (26)
Certainty	11.85 (26)	11.65 (26)	11.73 (26)	10.81 (26)
Diagnosticity	8.50 (26)	6.38 (26)	7.85 (26)	7.27 (26)

Note. Numbers in parentheses are cell *ns*. Differences in *ns* between measures are due to failure of some participants to complete all measures.