Does Self-Perception Change Explain the Foot-in-the-Door Effect?

Donald R. Gorassini and James M. Olson
University of Western Ontario

Consent to perform a small favor increases a respondent's susceptibility to perform a relatively large favor. This phenomenon, known as the foot-in-the-door effect, is considered to result from induced self-perception changes: the respondent comes to feel helpful for doing the small favor and complies again later out of a desire to maintain the instilled self-view. This study did not find a link between self-perception changes and large-request compliance in 2 experiments, although manipulations successfully altered self-rated helpfulness. Specifically, self-rated helpfulness increased (in Experiments 1 & 2) if participants' consent to a small favor brought social approval, and the ratings decreased (in Experiment 2) when social feedback for the small favor contained consensus information (i.e., indicated everyone else was also doing the favor). However, the results failed to predict either foot-in-the-door effects actually observed or compliance generally. Preexperimental gender differences in self-perceived helpfulness, in which women construed themselves to be more helpful than men, did successfully predict compliance with the large request. Implications for a theory of foot-in-the-door are discussed.

Several techniques for enhancing susceptibility to influence are used in social interaction (see Cialdini, 1993). One of them, aptly termed foot-in-the-door (FITD; Freedman & Fraser, 1966), consists of asking a person to perform a small favor. Presumably, compliance with the small request will enhance the probability of the person complying with a relatively large request made later. The FITD effect is demonstrated when compliance with the demanding request occurs at a higher rate or to a greater degree after the FITD tactic has been used (FITD condition) than when no susceptibility enhancement technique has been used (baseline). The FITD effect has been shown to occur at greater than chance rates across experiments, although the effect is usually small (for reviews, see Beaman, Cole, Preston, Kientz, & Steblay, 1983; DeJong, 1979; Dillard, Hunter, & Burgoon, 1984; Fern, Monroe, & Avila, 1986). Occasionally, large effects (e.g., Freedman & Fraser, 1966) have been observed, as have cases in which FITD produced less compliance than baseline (Brownstein & Katzev, 1985; Wang, Brownstein, & Katzev, 1989).

The dominant—indeed, almost universal—explanation of the FITD phenomenon has been a self-perception account. According to this view, doing someone a small favor stimulates a correspondent self-inference—the favor-doer comes to view him- or herself as helpful, or perhaps as compliant—and this self-perception change is responsible for the increase in further compliance (Beaman et al., 1983; Beaman, Svanum, Manlove, & Hampton, 1974; Cialdini, 1993; DeJong, 1979; Freedman & Fraser, 1966; Pliner, Hart, Kohl, & Saari, 1974; Snyder & Cunningham, 1975; Urbanowicz, 1975; Zuckerman, Lazzaro, & Waldgeir, 1979). Self-perception theory, in addition to providing this account of the FITD effect, furnishes a clue concerning why the FITD effect sometimes fails to occur: Compliance with the small request, though necessary, is not sufficient for enhanced compliance with the large request. Also necessary—and presumably sufficient—is self-perception change. The most important determinant of whether self-perception change will occur is thought to be the context under which the small favor is carried out. Specifically, if compliance is perceived by the actor as freely chosen (e.g., Zuckerman et al., 1979), as exercising more than a trivial personal cost (e.g., Seligman, Bush, & Kirsh, 1976), and as bringing social approval (Crano & Sivacek, 1982; Deutsch & Lambert, 1986; Goldstein, Seever, & Seever, 1982; Kraut, 1973; Moss & Page, 1972; Stimpson & Warranuntukle, 1988), then self-perception change will be more likely to occur than if some or all of these factors are absent. Thus, a possible reason why FITD effects have not always occurred in past studies is that these conditions have not always been present.

The findings of research assessing the self-perception account of FITD have been ambiguous enough to lead different reviewers, after referring to approximately the same literature, to draw

Preparation of this article was supported by Social Sciences and Humanities Research Council of Canada Research Grants. We thank Lynne Jackson, Susan Majewich, and Rudy Rab for their assistance in conducting the studies and Lanna Trick for her technical help with Experiment 1.

Correspondence concerning this article should be addressed to Donald R. Gorassini, Department of Psychology, King's College, 266 Euphorm Avenue, London, Ontario, Canada N6A 2M3, or to James M. Olson, Department of Psychology, University of Western Ontario, London, Ontario, Canada N6A 5C2.

\(^1\) The self-perception account is more a heuristic formulation than a complete theory of FITD. For example, it is not clear in current formulations how self-perception changes affect subsequent behavior. The parent theory (Bem, 1972) deals with how self-attribution is affected by behavior, but not with how self-views, once changed, affect further actions (Crano & Sivacek, 1982; Eisenberg, Cialdini, McCreath, & Shell, 1987). An alternative explanation of the FITD effect will be considered in the discussion section.
conclusions ranging from cautious support for the self-perception explanation (Beaman et al., 1983; DeJong, 1979) to the position that theoretical elaboration is necessary because some data do not fit easily into a self-perception framework (Dillard et al., 1984). We argue that drawing any firm conclusions from past research is difficult because direct measurement of self-perception change, as a way of validating pertinent experimental manipulations, has almost universally not been carried out (cf. Crano & Sivacek, 1982). To illustrate the problems of interpretation that arise, consider a hypothetical study in which the FITD effect is found when perceived choice is high but not when it is low (as the self-perception account would predict).

In the absence of direct measures, it would be impossible to tell whether the postulated self-attribution changes actually occurred. By the same token, if compliance rates were unaffected by a choice manipulation, it would be impossible to determine if this happened because no self-perception changes occurred or despite the fact that they did occur.

Preferably, then, the design of an experiment testing the self-perception account should include (a) FITD and no-FITD conditions, (b) self-perception measures taken after the small request but before the demanding one, and (c) appropriate compliance measures. To our knowledge, only one previous investigation meets these requirements. Rittle (1981) induced some undergraduate women to help an 8-year-old boy obtain candy from a malfunctioning machine. Compared to control women who were not exposed to this small opportunity to help, the experimental women subsequently volunteered more time to help with a research project. There was no difference, however, between experimental and control women's answers to a self-rating of helpfulness taken between the two requests for help, and self-ratings did not correlate with participants' behavioral responses to the second, large request. Although these null results raise questions about the self-perception account of the FITD effect, the single-item measure of helpfulness may not have been sensitive enough to detect possible self-perception differences.

In two other FITD experiments (Eisenberg et al., 1987; Scott, 1977), investigators also failed to find a relationship between self-rated helpfulness and large-request compliance. However, the self-ratings were elicited after the large request, so they may not have reflected the self-views that were actually salient in the crucial period between the small and large requests.

Thus, those researchers who have included self-perception measures in their FITD studies have obtained rather disappointing results. Stronger support for the self-perception analysis has come from studies of prosocial behavior where investigators have assessed the influence of compliance with an initial request on self-perceived helpfulness; second requests were not made in these studies. In two experiments by Williamson and Clark (1989), virtually all participants who were asked to help gather materials that the experimenter "accidentally" knocked to the floor agreed to do so, and all who agreed were thanked and called helpful. After they gave their help, these participants made higher self-ratings of helpfulness than did control participants who received no request. In other investigations, self-perceptions of helpfulness have been shown to depend on whether doing the favor is perceived by the helper as freely chosen. For example, when compliance occurred under high external justification (low choice), such as after a promise of payment (Batson, Coke, Janoski, & Hanson, 1978; O'Malley & Andrews, 1983; Smith, Gelfand, Hartmann, & Partlow, 1979), after participants were required to read about persons whose altruism was exemplary (Thomas, Batson, & Coke, 1981), or in the context of an exchange (as contrasted to a communal) relationship (Williamson & Clark, 1989, Experiment 3), self-ratings of helpfulness were lower than those made under conditions of low justification (high choice). Taken together, these studies strongly support the half of the self-perception account that pertains to how compliance with the initial request influences ensuing self-ratings of helpfulness.

Our purpose in two investigations was to test the self-perception account of the FITD effect more directly and more fully than in past studies. First, we manipulated factors thought to invigorate the self-perception process and result in self-attributions that would promote compliance with the large request. Second, as did Rittle (1981), we measured self-attributions formed in the interval between the small and large requests, but our measures were more extensive than in the earlier study. In the assessment, we presented several trait words and asked participants to respond yes to terms considered true of the self and no to terms considered untrue. The direction of participants' answers (i.e., yes vs. no) provided one measure of self-perception. In addition, the latency of these responses was measured. We deemed latency a reflection of whether self-labeling had occurred in the participant's mind prior to the self-rating task. Presumably, for participants already thinking of themselves as helpful—as might happen after consenting to or actually performing a small favor—little time would be required for a yes response to such terms as helpful or generous. In effect, answers would be formulated beforehand. A participant responding slowly to such terms would be someone without a ready answer. In the absence of an answer, the time necessary for a decision would be taken after presentation of the trait term (see Fazio, 1989).

Theorists have noted that the precise nature of the self-label induced by compliance with a small request is unclear (see Crano & Sivacek, 1982; DeJong, 1979). It has been proposed that consent could change self-perceptions of helpfulness or of submissiveness, either of which could have an impact on subsequent compliance (Batson, Harris, McCaul, Davis, & Schmidt, 1979; DeJong, 1979). To examine this issue, we included terms related to each trait in our self-rating task. We were also able to assess the rapidity of self-perception change. Beaman et al. (1983), in their meta-analysis, distinguished between studies in which the time separating the small and large requests was short (less than 1 day) and studies in which it was long (1 day or more) and found that the FITD effects produced in the two cases were of approximately equal magnitude. However, the correlation between small-request compliance—an indirect measure of self-perception change—and large-request compliance was unexpectedly negative in trend over short-delay studies but positive in trend over long-delay studies. It was suggested, consequently, that a different process might underlie the FITD effect in short-delay as compared to long-delay research (Beaman et al., 1983) and that "if
the second request is made soon after the initial request, the cognitive work necessary for a self-perception change may not or perhaps cannot take place (Beaman, Steblay, Preston, & Klentz, 1988, p. 234). Research conducted since, in which delay and self-perception were manipulated but putative self-perception effects were not measured, failed to provide definitive evidence either for or against the delayed self-perception change hypothesis (Beaman et al., 1988).

It is noteworthy that self-perception theory and research would not otherwise spur investigators to anticipate long delays in self-perception change onset. Such alteration is routinely produced in a single experimental session (e.g., Bem, 1972; Olson, 1992; helping research reviewed above). In the two experiments reported in this article, our small-request, self-perception measures, and large request all occurred in one session lasting less than 1 hr. By taking intervening self-perception measurements soon after the small-request phase, we were in a position to determine whether or not the onset of self-perception change can be rapid.

**Experiment 1**

The design of our first study included three conditions, two in which a FITD tactic was present and one in which it was absent (baseline). The sequence of events in all cases was introductory instructions first, self-report tasks next, and large request last. In the FITD conditions, the experimenter included a request for a small favor in her introductory discourse, and she responded to compliance with either extended positive feedback (extended-thanks condition) or a thank you (thanks-alone condition). The responses she provided were chosen to represent the opposite extremes of the range of feedback types provided in actual FITD situations (see Crano & Sivacek, 1982) and in small-favor situations generally. The self-perception task required participants to rate numerous computer-administered trait terms for self-descriptiveness. Finally, participants were exposed to a relatively demanding request.

**Hypotheses**

The design made possible a test of several hypotheses derived from the self-perception account of FITD. Our small request, which was not trivial in its consequences and which participants were reasonably free to refuse, possessed features capable of activating and directing self-perceptions. With respect to activation, we assumed that participants in our baseline group would not actively evaluate their degree of helpfulness prior to making the computer self-ratings. Therefore, their response times could be interpreted as baseline latencies, that is, as times taken by individuals who have not been given an opportunity prior to the computer test to draw conclusions about their helpfulness and who, consequently, would have made the evaluations on the spot. Faster response times than baseline, as exhibited in any of the study’s two experimental conditions, would suggest that participants had already been thinking about their level of helpfulness and had settled on a particular label before the computer task. Accordingly, we could test two competing accounts of how readily the self-perception process is activated. Crano and Sivacek (1982), who have written specifically on this issue, argued that high levels of social approval are necessary for activation. This reasoning would predict faster than baseline responding to help-related terms in the extended-thanks condition only. The alternative account, implicit in FITD research, assumes that the customary social feedback for consent to a favor (i.e., a thank you) is sufficient for activation. This analysis leads to the prediction of faster than baseline responding in the thanks-alone condition as well as in the extended-thanks condition.

With respect to the direction and extent of self-perception change, we expected participants to adjust their self-perceptions according to observations of their own behavior and the context (cf. Bem, 1972; Olson, 1992; Olson & Zanna, 1993). After helping someone, and especially after helping someone and receiving pronounced social approval, participants would have lucid evidence of their readiness to render aid. Hence we predicted that higher-than-baseline ratings of helpfulness would occur in the two FITD conditions, and the highest ratings of all would arise in the extended-thanks group.

Compliance with a large request is assumed in the self-perception account of FITD to parallel the level of self-perceived helpfulness produced by compliance with the small request. This led to the prediction of a straightforward relationship between induced self-perception changes and actual helpfulness: The greater the increase in self-perceived helpfulness, the greater the increase in compliance with the large request.

**Method**

**Participants.** A total of 140 introductory psychology students at the University of Western Ontario, 72 women (mean age = 19.51) and 68 men (mean age = 20.03), participated for course credit in a 1-hr session entitled “Computer Personality Assessment.” Participants were randomly assigned to one of the three conditions. Dropped from the analyses were 5 participants (1 woman and 4 men)—2 for failing to complete all questionnaire items and 3 because they suspected the large request was part of the experiment. No suspicion was expressed by anyone about the small request.

**Procedure.** Each participant was tested individually in a small room by a female experimenter. A reminder that the session would last exactly 1 hr was given at the outset. Next, a compliance induction consisting of a small request and a prearranged form of response-contingent feedback was either provided (in FITD conditions) or not provided (baseline). The following request was the FITD tactic used:

> There is something I’d like to ask of you. It has nothing to do with this study, or your credit; it’s a favor. I’m working on preparing another study right now and am trying to get the materials together. One thing I have to do is rate some short written paragraphs for clarity and style. It’s a simple task, but as it turns out, I can’t use just my own rating; I need to compare two people’s ratings. Could you stay for five, and at most ten, minutes at the end of the experiment to help me with that?  

\(^2\) During the academic term in which this study was run, most classes commenced on the half-hour and usually were of 1-hr duration. Given this, experimental sessions were scheduled on the hour to ensure that participants did not have classes beginning immediately after the study (which would have forced them to decline the small request). Staying for 10 min after the session, which the small request asked participants to do, therefore required the donation of personal time.
In the extended-thanks condition, if the participant agreed to stay, the experimenter stated: “Thanks, that’s very generous of you. I know students have very little time for extra work. I really do appreciate you doing this for me, and believe me it helps out a lot.” By contrast, in the thanks-alone condition, the experimenter responded “thank you” if the participant agreed to stay. In either of the FITD conditions, a refusal of the request drew the response: “Okay, thanks anyway.” Five participants in the extended-thanks condition and three participants in the thanks-alone condition refused the small request (but were retained in the analyses; see DeJong, 1979).

After these apparent preliminaries, the cover story was conveyed in all conditions:

As you know, we are interested in examining personality with the help of a computer. Traditionally, personality assessments have been carried out using paper-and-pencil methods. But since computers are being employed more and more for this purpose, and because surprisingly little is known about how they influence the measurement process, we feel it is important to compare the traditional and the modern approaches in a controlled fashion. Therefore, you will be administered several personality items both on the computer and in paper-and-pencil format.

Self-rating tasks were assigned next. In the first of these, several trait terms were presented in sequence by computer, and the participant judged each word for its self-descriptiveness. Information on the computer screen introduced the task and instructed participants (a) to respond to each word by pressing the key labeled yes (always 0 on the main keyboard) if the term was self-descriptive, or by pressing the key labeled no (always 1 on the main keyboard) if it was not, and (b) to make the response accurately but also as quickly as possible (for guidelines, see Fazio, 1989). Participants were required to keep prepared by placing their index fingers over the “yes” and “no” keys.

After going through the instructions and eight practice trials under the experimenter’s supervision, participants responded without supervision to 32 randomly ordered critical adjectives: 8 related to helpfulness, 8 to submissiveness, and 16 not related to either helpfulness or submissiveness. A given trial, practice or critical, consisted of a 2-s display of 14 asterisks, aligned in a row in the screen’s center, followed by the immediate replacement of the asterisks by a trait term that remained on the screen until the participant’s key press response was made. The key press immediately brought on the next trial. Response (yes vs. no) and response latency, to the millisecond, were automatically recorded on each trial. Participants were not informed that responses were being timed.

Following the computer task, paper-and-pencil versions of the Self-Monitoring Scale (Gangestad & Snyder, 1985) and Role-Play Scale (Fletcher & Averill, 1984) were completed. These scales were described as being representative of traditional personality measures.

Next, a 5-min intermission in which the experimenter left the room, purportedly to print out test results and prepare for the remainder of the session, was provided. Before leaving, she presented the following demanding request, stating:

Now that there is some free time, you might as well read this. It is a departmental memo being circulated to all first year research participants and concerns a problem in the department. So much research is going on that there are not enough students in the subject pool to fill the demand. As a consequence, some of our graduate students cannot complete their research because there are too few subjects. The department is therefore asking first year students to consider volunteering extra research time—it is in addition to credit time—to help the grad students out. The memo explains all this. Please read it while I’m gone. If you wish to volunteer extra time, you can fill out this bottom portion.

The memo recapitulated this information. For those deciding to volunteer, spaces for name and date were provided. In addition, the alternatives 1 hour, 2 hours, 3 hours, and other were listed. Pledges were made by circling an existing number or by writing a number in a space next to “other.” When the experimenter returned, participants were probed for suspicion and debriefed.

Measures. Two self-rating scores, one for helpfulness and one for submissiveness, were computed per participant on the basis of yes or no responses made during the computer task. First, all negatively keyed terms (e.g., selfish) were reverse scored. A helpfulness self-rating was the percentage, over the eight helpfulness trait terms, in which the self was rated as helpful. A submissiveness self-rating was the percentage, over the eight submissiveness terms, in which the self was rated as submissive. Three latency scores were computed per participant by separately averaging response times to the eight helpfulness adjectives, the eight submissiveness adjectives, and the 16 neutral terms. Each set of latency scores was positively skewed and therefore transformed for analysis using log10. Compliance with the large request was expressed as the number of hours volunteered for experimental participation.

Results

Preliminary analyses that included gender as a variable were conducted. This examination revealed that women rated themselves as more helpful, $M = 92.39$ versus 85.38, $F(1, 129) = 5.56, p < .02$, and more submissive, $M = 42.25$ versus 30.25, $F(1, 129) = 10.56, p < .001$, than did men; and women made their helpfulness ratings faster than did men, untransformed $M = 1,320$ versus 1,557 ms, $F(1, 129) = 8.46, p < .01$. Women also backed their words with deeds: They volunteered more hours in response to the large request than did men, $M = 1.06$ versus 0.67, $F(1, 129) = 4.44, p < .04$. There were no interactions, however, between sex of participant and conditions on any dependent measure, so gender did not qualify the predictions tested below.\(^3\)

Helpfulness terms were: unhelpful, self-centered, selfish, hard-hearted, generous, charitable, helpful, and giving. Submissiveness terms were: unyielding, strong, independent, nonconforming, submissive, conforming, follower, and compliant. Neutral terms were: superstitious, reckless, unproductive, noisy, sloppy, scatterbrained, athletic, methodical, healthy, intelligent, careful, tidy, muscular, prompt, slowpoke, and simple. Pretesting indicated that correlations among ratings of terms within the helping category and within the compliance category were high and significant, whereas correlations between ratings of words from different categories—helping, compliance, or neutral—were low and generally nonsignificant.\(^4\)

Although they were included primarily to protect the cover story, we thought that self-monitoring and/or role-playing ability might influence the self-perception process. Exploratory analyses revealed no interesting or easily interpretable effects for these variables on the self-rating and compliance measures, however. Thus, these individual difference dimensions will not be discussed further.\(^5\)

It may also be worth noting that these preliminary ANOVAs revealed significant ($p < .05$) main effects for the conditions factor on all measures that yielded differences in the planned contrasts reported in the following sections, with one exception: The conditions main effect did not reach conventional levels of significance on the helpfulness trait measure, even though this measure confirmed one of our planned contrasts.
We used planned comparisons in analyzing data pertinent to our central hypotheses concerning self-perception and compliance. This consisted of contrasting (a) the FITD groups, as a unit, with the control group and (b) the extended-thanks group with the other two groups.

**Self as helpful.** Untransformed mean response times, in ms, in the baseline, extended-thanks, and thanks-alone conditions were 1,568, 1,384, and 1,340 respectively. The two FITD conditions combined fell below baseline, $t(132) = 3.18, p < .01$. The contrast between the extended-thanks group and the other two groups combined was not reliable ($t < 1$). Mean self-ratings in the baseline, extended-thanks, and thanks-alone groups were 85%, 92%, and 91%, respectively. The two FITD conditions combined exceeded baseline, $t(132) = 2.30, p < .02$. However, helpfulness ratings were no higher in the extended-thanks group than in the other two groups ($t < 1$). To examine the possibility that response times determined self-ratings, we also used analysis of covariance (ANCOVA) to remove variance due to response times from the self-rating data. This did not affect the above reported self-rating findings.

To summarize, among participants who received the small request, whether feedback for consent to that request consisted of extended thanks or thanks alone, (a) the self-perception process was activated, and (b) introspection resulted in the self being labeled as relatively helpful.

**Other self-ratings.** Similar analyses were applied to self-perceived submissiveness data as were applied to the self-perceived helpfulness data. None of the planned comparisons involving response times or self-ratings was reliable at or near $\alpha = .05$. Neutral scores also failed to vary over conditions.

**Compliance with the large request.** As recommended by DeJong (1979), compliance data of all participants, even those who refused to perform the small favor, were included in the analysis (as was the case in all preceding analyses). Mean number of hours volunteered in the baseline, extended-thanks, and thanks-alone conditions were 0.87, 1.16, and 0.60, respectively. The two FITD conditions combined did not differ from the control group, $t < 1$. However, more hours were pledged in the extended-thanks condition than in the other two groups combined, $t(132) = 2.31, p < .02$. When the data of the eight participants who refused the small request were excluded from the analysis, the findings were not altered.

**Correlations.** The relationship between each self-perception index (helpfulness and submissiveness ratings; helpfulness, submissiveness, and neutral latencies) and number of hours volunteered was examined. None of the correlations, either overall or within-cell, reliably exceeded zero. In view of the gender differences in self-ratings and in compliance that were observed, we repeated the aforementioned correlational analyses for each gender separately. One correlation, that between compliance and self-rated helpfulness, was reliable in the case of women, $r(69) = .34, p < .05$, but not men, $r(62) = -.04, n.s.$

Further attempts were made to detect a possible relationship between self-perceived helpfulness and compliance. Alternative 1 consisted of removing variance attributable to self-reporting in favorable fashion from helpfulness self-ratings and subsequently examining correlations between the resulting "pure" helpfulness index and compliance. Accordingly, neutral traits were scored for self-reporting in a favorable manner (e.g., unproductive was negatively keyed and intelligent positively keyed), the scores were summed to form an approval index, and partial correlations between helpfulness, with approval removed, and compliance were computed overall and within experimental conditions. None of these partial correlations was reliable.

Alternative 2 consisted of combining helpfulness ratings and latencies into a single self-perception index and computing correlations between this composite and compliance. The method for combining the helpfulness variables followed our conception of what the measures were tapping: self-ratings reflected direction and extremity whereas latencies reflected activation. An extreme and activated self-label, we assumed, should exert more impact on behavior than either a moderate or inactive self-label. Thus an extreme and active "helpful" label should be associated with the highest levels of compliance, and an extreme and active "unhelpful" label should be associated with the highest levels of refusal (i.e., lowest levels of compliance). Accordingly, helpfulness self-ratings and latencies were each standardized. Latency standard scores were further transformed by setting $z$ scores of greater than 1 equal to 1 (low activation), $z$ scores in the interval 1 to −1 equal to 2 (moderate activation), and $z$ scores less than −1 equal to 3 (high activation). A composite self-perception score was computed by multiplying the participant’s standardized helpfulness self-rating by his or her activation coefficient of 1, 2, or 3. As above, this composite score also failed to correlate with compliance either overall or within cells. In view of the fact that a majority of participants’ self-ratings were in the direction of greater helpfulness, we computed a second, simpler self-rating–latency composite in which each participant’s standardized helpfulness self-rating and her or his reverse-scored standardized helpfulness latency were summed. This summed composite also failed to correlate with compliance either overall or within cells.

**Discussion**

The measures of self-perception and compliance yielded different patterns of findings. We will discuss the results for each type of measure individually and then consider their relationship.

**Self-perception change.** Ratings of helpfulness were faster and higher than baseline in the extended-thanks and thanks-alone groups. These findings are consistent with the notion that the small-request situation rapidly got participants thinking about how helpful they were and subsequently directed them to the conclusion that they possessed a relative abundance of the trait. Our data confirm previous findings indicating that recipients label themselves as helpful soon after consenting to a favor, especially when the favor is not trivial and when consent is freely chosen (e.g., Batson et al., 1978; Smith et al., 1979; Williamson & Clark, 1989). Our measures of self-perception were also more extensive than in past studies; indeed, we know of no previous demonstrations in which exposure to a small request produced faster self-ratings of helpfulness. We believe that this is important because it establishes that, prior to the self-rating
task, participants reflected on and drew conclusions about their helpfulness.

With respect to this self-perception activation, response times to help-related terms were as fast when agreement to perform the small favor received thanks alone as when the agreement brought extended thanks (and, recall, the responses were faster than baseline in both conditions). This pattern contradicts the notion that extended thanks is necessary for self-perception activation (cf. Crano & Sivacek, 1982).

It is noteworthy that, once a self-perception process is activated, a thank you can be as effective at enhancing ratings of helpfulness as feedback consisting of a thank you, much praise, and the label generous. It should be noted, however, that all participants rated themselves as quite helpful. Even in the control group, an average of 6.8 helping words out of 8 were endorsed; in the extended-thanks and thanks-alone conditions, the rates were 7.4 and 7.3 of 8, respectively. Perhaps extended thanks would have produced higher self-rated helping than thanks alone if there had been more room for upward movement along the scale. We will return to this issue in Experiment 2.

The FITD literature typically refers to helpfulness as the disposition that individuals attribute to themselves for agreeing to perform the small request; submissiveness or compliance is the other trait sometimes mentioned as a possibility (DeJong, 1979). To our knowledge, the present study is the first to investigate these alternatives concurrently. The data were clear: Helpfulness, not submissiveness, was used by these participants to label their performance of a favor. As noted above, the content of the experimenter's small request defined the situation as helping. Participants likely accepted this interpretation. It is likely, too, that the tendency for humans to label themselves in the most positive terms available (see Greenwald, 1980; Olson & Hafer, 1990; Taylor & Brown, 1988) also influenced the self-inference process. Of course, it is more flattering to think of ourselves as helpful and generous than as weak and submissive.

Although we did not make any predictions about gender differences, preliminary analyses showed that, across conditions, women rated themselves as more helpful and more submissive than did men, and women made their helpfulness ratings faster than men. Nurturance, compassion, and submissiveness are part of the female gender role (Bern, 1974; Ruble, & Hafer, 1990; Taylor & Brown, 1988) also influenced the self-inference process. Of course, it is more flattering to think of ourselves as helpful and generous than as weak and submissive.

Compliance. The extended-thanks condition produced a higher rate of compliance with the large request than did the other two conditions. Thus, extended social approval was necessary for the FITD effect to occur. This result confirms the data of several previous investigators who found that thanks alone did not produce greater compliance than control conditions (Crano & Sivacek, 1982; Stimpson & Waranusuntikule, 1988). These researchers generally interpreted their results within a self-perception framework, arguing that social approval is necessary for the FITD effect because it activates and directs self-reflections about helpfulness. As we will discuss in the next section, however, there are problems with applying such an interpretation to our data.

Consistent with the obtained sex differences on the self-perception measures, the compliance measure showed that, across conditions, women volunteered more time in response to the large request than did men. This finding seems compatible with the social-role perspective on helping proposed by Eagly and Crowley (1986), who argued that gender roles encourage different kinds of helping by men and women. Specifically, men are more likely than women to exhibit helping that is heroic and chivalrous, whereas women are more likely than men to exhibit helping that is caring and nurturant. The preponderance of past social psychological studies of helping have involved male-oriented responses (e.g., helping a victim in distress), and the typical finding has been greater helping by men (see Eagly & Crowley, 1986, for a meta-analysis). But some researchers have found differences in the opposite direction, and the settings in these studies have usually been more interpersonally nurturant—including volunteering to be in a later experiment (e.g., Bickman, 1974, Experiment 3). The large request in the present experiment explicitly appealed to participants to "help the experimenter out"—an appeal that seems more relevant to a nurturant than an heroic motive. At any rate, our finding adds to the growing evidence that men do not always help more than women; in fact, the opposite is often true.

Self-perceptions and compliance. If the compliance data of the present investigation are considered in isolation, as if self-ratings had not been elicited, support for the self-perception account might be seen in the findings. The fact that compliance levels were highest in the condition designed to maximize self-perception change squares with such an account, and the less than expected amount of compliance in the thanks-alone condition could be attributed to the inability of a "thank you" to stimulate a self-labeling process. In any event, the compliance data on their own provide no serious threat to the self-perception account.

When the self-rating data are juxtaposed with the compliance findings, however, problems for the self-perception account become apparent. Specifically, consistent with previous studies that have included both self-rating and behavioral measures (Eisenberg et al., 1987; Rittle, 1981), the expected parallel between the self-perception and compliance measures failed to occur. Whereas the compliance measure revealed a FITD effect only in the extended-thanks condition, the self-perception measures showed increased perceptions of helpfulness in both the extended-thanks and the thanks-alone groups. Thus, participants in the thanks-alone condition made higher than baseline ratings of helpfulness but did not exhibit higher than baseline compliance with the large request.

Even though the experimentally induced changes in self-perception failed to bring about changes in compliance, preexperimental gender differences in self-perceptions did parallel compliance. Women's ratings of helpfulness were higher (and faster) overall than were those of men, women complied more overall than did men, and the correlation between helpfulness and compliance was significant for women but not for men. These data, taken together, suggest that compliance will be affected by self-perceptions of a long-standing sort but be unaffected by self-rating changes that result from brief, single-trial tactics like FITD. Indeed, some of the most dramatic compliance effects appear to occur when a series of requests for favors of gradually increasing magnitude are applied (cf. Cialdini, 1993).
In fact, one could argue that the gender differences provide evidence of convergent validity for our self-perception and compliance measures. By showing that parallel effects are possible, these data add credence to our conclusion that the between-conditions divergence of the measures indicates that the impact of the manipulations on compliance (i.e., the FITD effect) was not mediated by self-perception changes.

**Experiment 2**

In a second experiment, we manipulated two separate factors that the self-perception account suggests will affect compliance in the FITD situation. One of the factors was social approval, which is what we examined in Experiment 1. The second factor consisted of either telling or not telling participants just after they agreed to the small request—which, as in Experiment 1, was to help for a few minutes after the session—that almost every other participant was also agreeing to stay. If, in helping someone, one learns that everyone else readily helped as well (i.e., if there is high consensus), then the rendering of aid is likely to be attributed to the situation (cf. Kelley, 1973). If helping happens to be in response to the initial small request of a FITD sequence and occurs under high-consensus information, then attributions to a helpful disposition are unlikely. This movement away from a dispositional attribution and toward a situational one should, according to the self-perception view, be associated with decreased levels of large-request compliance.

DeJong (1981) assessed the role played by consensus in the FITD situation in a study in which mall patrons were asked to sign a petition in support of an organization for people with physical disabilities. After consent to the favor occurred, one of two types of feedback was delivered. Participants were either shown a sheet almost full of names and told everyone who was approached had signed (high consensus) or were shown a sheet with no names on it and told no one who was approached had signed (low consensus). Contrary to self-perception theory predictions, participants in the high-consensus condition later helped (a confederate who dropped a quarter) at a higher than baseline rate whereas participants in the low-consensus condition helped at the baseline level. These data are interesting because they raise the possibility that FITD effects can occur in the absence of the subjective changes defined by self-perception theory. However, self-perception measures were not taken, so whether self-perceived and behavioral helpfulness actually failed to correspond cannot be ascertained.

In testing the respective roles of social approval and consensus information, we ran a study in which the small request, self-perception measures, and large request of Experiment 1 were used again. This time, however, four (instead of two) different feedback conditions were included in the design. Participants were either shown extended gratitude or were given thanks alone for agreeing to the small request. Also, following DeJong (1981), participants were either told or not told that everyone else was agreeing to the favor. The factorial combination of the two social approval levels (extended feedback vs. thanks alone) and the two consensus information levels (present vs. absent) made for four different feedback conditions. A no-small-request condition was also included in the design as a baseline.

The design also included three other new features. First, we actively sought to ensure the anonymity of responses to the large request. In the self-perception account of FITD, the desire to comply with the large request is a private concern. By making responding to the large request anonymous, we eliminated the experimenter as a source of compliance pressure, leaving in the wake of this outside incentive the participant's own self-definitional concerns, for example, knowledge of being a helpful person. The self-perception account would lead us to expect a FITD effect even under these private conditions. Second, mood was measured. Mood can be instrumental in helping behavior (e.g., Cunningham, Shaffer, Barbee, Wolff, & Kelley, 1990). Perhaps mood changes, as produced by FITD manipulations, are at the root of compliance changes. Third, a measure of expectations specific to the large request were assessed. Perhaps FITD manipulations, before all else, change expectations about the effects of responding to appeals like the experimenter's large request, in this case participation in further experiments. These expectations, in turn, might affect compliance.

**Hypotheses**

**Self-perception.** The following hypotheses pertain to the effects of our manipulations on self-perceived helpfulness. We did not expect the manipulations to have any effect on self-perceived submissiveness.

One set of predictions within this category concerns only that portion of our design represented in the 2 X 2 (social approval X consensus) factorial combination. First let us address expected effects for response latencies. We could test again the two accounts of how social approval affects activation of the self-perception process. If high levels of approval are necessary, then faster response times to help-related words would be observed when feedback contained extended thanks than when it contained thanks alone. If the customary response to a small favor (i.e., a thank you) is sufficient to trigger the process, as in Experiment 1, then the two conditions would not necessarily differ. More will be said about this issue below, where we discuss differences in expected activation levels between baseline and the two extended-thanks conditions.

With respect to the consensus level manipulation, we assumed that exposure to consensus information would provoke a conflict between an existing, desirable self-conception, that is, caring and helpful, and the opposite view implied in the consensus information, that is, weak and ingratiating. This conflict, instead of speeding up responses to help-related words, would likely keep them slow. We expected, therefore, to see slower response times when consensus information was provided than when it was not provided.

Concerning the direction and extent of self-perception changes, we could again test the hypothesis—not confirmed in Experiment 1—that extended feedback would result in higher self-rated helpfulness than would thanks alone. For the first time we could also test the role of consensus level in rated helpfulness. On the assumption that consensus information results in the self being discounted as a source of apparently helpful behavior, we expected to observe lower ratings of helpfulness in the consensus condition than in the no-consensus condition.
A further set of predictions about the self-perception process concerned each of the four small-request conditions in relation to baseline. Predictions about response times were based on the following reasoning: In the two conditions in which consensus information was provided, a conflict between the self as caring and helpful and the self as weak and ingratiating was expected to slow down responses to help-related words. In the control condition, a self-perception process should not occur and, consequently, response times would be relatively slow. A clear prediction to arise from the foregoing is that response times in the extended-thanks–no-consensus condition would exceed baseline. Moreover, if thanks alone is sufficient to activate the self-perception process—as it appears to be (see Experiment 1)—then faster than baseline responses to help-related terms would also occur in the thanks-alone–no-consensus condition. Neither of the two consensus information conditions (extended-thanks–consensus nor thanks-alone–consensus) was expected to produce faster than baseline responses because consensus information would contradict the notion of helpfulness.

In understanding our predictions concerning the direction and extent of self-perception changes, consider that one of our manipulations, social approval, was designed to augment self-perceived helpfulness levels whereas the other manipulation, consensus level, was designed to reduce those levels. The greatest increase relative to baseline in self-ratings, therefore, was expected to occur in the condition with the augmenting factor present and the reducing factor absent, that is, extended-thanks–no-consensus. By contrast, the greatest decrease relative to baseline in self-ratings was expected to occur in the condition with the reducing factor present and the augmenting factor absent, that is, thanks-alone–consensus. Regarding the other two small-request conditions: if thanks alone is an augmenting factor, as it appeared to be in Experiment 1, then an increase above baseline in helpfulness ratings would be expected in the thanks-alone–no-consensus condition; and, in that extended feedback and consensus information are antagonistic to one another in their effect on self-perceived helpfulness, no departure from baseline helpfulness ratings was expected with the presentation of both factors (i.e., in the extended-thanks–consensus condition).

Compliance. Self-perception theory suggests a straightforward relationship between self-perceptions and compliance with a large request: the more helpful one considers her- or himself to be, the greater the compliance that will occur. Thus, those conditions that produced the highest and fastest ratings of helpfulness were expected to produce the largest rates of compliance. Also, ratings of helpfulness were expected to correlate with levels of compliance.

Method

Participants. A total of 185 introductory psychology students at the University of Western Ontario, 86 women (mean age = 19.55) and 99 men (mean age = 19.73), recruited as in Experiment 1, were randomly assigned to the 5 conditions of a $2 \times 2$ (social approval $\times$ consensus information) factorial with added no-small-request baseline group design and run by one of two experimenters, a man or a woman. Perhaps due to our procedures for ensuring the appearance of anonymity of large request responses, no participant registered suspicion about the study’s purpose.

Feedback conditions. As noted, the small request, the computer assessment of self-perceptions, and the large request were the same here as in Experiment 1. The extended-thanks–no-consensus, thanks-alone–no-consensus, and control conditions of Experiment 2 consisted of the same feedback as the extended-thanks, thanks-alone, and control conditions, respectively, of Experiment 1. The two new conditions of Experiment 2 involved the addition of consensus feedback to extended-thanks feedback as well as to thanks-alone feedback. To summarize, then, feedback in the present study for consent to the small favor consisted of either (a) extended thanks, (b) a thank you, (c) extended thanks followed immediately by consensus information, or (d) a thank you followed immediately by consensus information. Consensus information consisted of the following:

- Actually, almost everyone volunteers when I ask. Let’s see, in the week-and-a-half since I’ve been asking for volunteers, we’ve run 22 people in this experiment, and all 22 have agreed to stay the extra ten minutes, which, psychologically, is interesting in itself. I guess 10 minutes isn’t a lot to ask, or it could be that it’s very hard to turn down a request of any type in a psychology experiment.

Mood measure. Just prior to the computer assessment of self-perceptions, which in experimental conditions was also just after participants received feedback for their response to the small request, a paper-and-pencil version of the Positive and Negative Affect Schedule was administered (PANAS, Watson, Clark, & Tellegen, 1988). Participants rated their current mood by responding to each of the test’s 20 adjectives—10 positive (e.g., enthusiastic or interested) and 10 negative (e.g., scared or upset)—using a 5-point scale (1 = very slightly or not at all through 5 = extremely). The test yields two scores—one for degree of positive affect (range 10–50) and one for degree of negative affect (range 10–50).

Anonymity of response to large request. Once the self-perception test was completed, as in Experiment 1, the experimenter ejected a diskette from the computer, got up, and informed participants she was leaving for at least 5 min to score and print test results. As an apparent afterthought, she indicated it would be a good time for the participant to read a memo from the Psychology Department. This was the same memo used in Experiment 1 to convey the large request and provide the participant with an opportunity to volunteer research time. Participants were asked to read the memo. They were also shown a sealed ballot box marked Psychology Department in which to put completed memos and a pile of blank memos on which to replace their own form if they did not complete it. From the participant’s point of view, the returning experimenter would not know whether a completed memo had been deposited into the ballot box or an uncompleted one had been put back onto the pile of blanks.

Expectation measure. When the experimenter returned, participants were administered a scale designed to measure their expectations regarding the immediate effects of further experimental participation. The test consisted of seven items with 7-point response scales, anchored as follows: rewarding–unrewarding, pleasant–unpleasant, tense–calm, pain-in-the-neck–not-a-pain-in-the-neck, friendly–unfriendly, good-time–not-a-good-time, worthwhile–not-worthwhile. After reverse scoring, the scale scores were summed to form a single index (range 7 to 49) with higher values representing more positive expected experiences. The scale’s internal consistency was high (coefficient $\alpha = .91$). After completing this scale, participants were probed for suspicion and debriefed.
Results

Preliminary analyses did not reveal any differential effect on dependent measures of the study's two experimenters. Therefore, the experimenter factor is not included in the analyses reported below.

Self-perceptions. Figure 1 gives the mean latencies (untransformed) to help-related terms, and Figure 2 the mean self-ratings of helpfulness, within each condition including baseline. We first analyzed the effects of the social approval and consensus level manipulations on self-rated helpfulness, without considering the no-small-request (i.e., baseline) group. Comparisons of experimental groups with baseline are reported below. A 2 X 2 X 2 (social approval x consensus level x sex) completely randomized analysis of variance (ANOVA) on transformed latencies to help-related terms indicated there were longer response times in the consensus condition \( (M = 1505) \) than in the no-consensus condition \( (M = 1359) \), \( F(1, 139) = 4.67, p < .05 \), and among men \( (M = 1555) \) than among women \( (M = 1295) \), \( F(1, 139) = 11.58, p < .001 \). Neither the social approval main effect nor any of the interactions were reliable. A similar analysis of helpfulness self-ratings yielded three reliable main effects: helpfulness was rated higher after extended thanks \( (M = 88.78) \) than after thanks alone \( (M = 82.43) \), \( F(1, 139) = 4.88, p < .05 \), lower after consensus information was provided \( (M = 82.24) \) than when withheld \( (M = 89.61) \), \( F(1, 139) = 7.17, p < .01 \), and higher among women \( (M = 90.41) \) than men \( (M = 81.69) \), \( F(1, 139) = 13.02, p < .001 \). No interactions were reliable.

Next we examined experimental condition-baseline contrasts. As predicted, (transformed) latencies were shorter than baseline in the extended-thanks-no-consensus condition, \( t(175) = 3.18, p < .001 \), and not unexpectedly, shorter than baseline in the thanks-alone-no-consensus group, \( t(175) = 2.64, p < .01 \). Response times in each of the two consensus conditions failed to deviate from baseline \( (ts < 1) \). Similar analyses were conducted on self-ratings of helpfulness. Also as expected, the ratings made in the extended-thanks-no-consensus group exceeded baseline (i.e., in the direction of greater helpfulness), \( t(175) = 1.67, p < .05 \), whereas ratings made in the thanks-alone-consensus group fell below baseline, \( t(175) = -1.71, p < .05 \). Also consistent with predictions, ratings made in the extended-thanks-consensus and thanks-alone-no-consensus conditions failed to deviate from baseline levels \( (both ts < 1) \).6

Submissiveness, mood, and expectations. Using analyses similar to those used to examine self-perceived helpfulness data, we found no influence of our manipulations on response times to submissiveness-related terms, or to self-rated submissiveness, mood, or expectation.

Compliance. The 2 X 2 X 2 (social approval x consensus level x sex) ANOVA on hours pledged yielded a significant three-way interaction, \( F(1, 139) = 4.66, p < .05 \), which is represented in Figure 3. Among women, approximately the same amount of time was pledged from one condition to the next. In the case of men, however, extended thanks and consensus level interacted, \( F(1, 75) = 10.51, p < .01 \), with more time pledged in the extended-thanks-consensus condition than in the three remaining conditions. Other than the three-way interaction, none of the effects in the analysis were reliable at \( \alpha = .05 \).

Next, time pledged in the control condition was compared to time pledged in each experimental condition. We carried out this task separately for men and women because a 2 X 5 (sex x condition) ANOVA yielded a reliable two-way interaction, \( F(4, \)

---

6 We also used analyses of covariance (ANCOVAs) in Experiment 2 to examine the possibility that response times determined self-ratings. The difference among groups in terms of self-rated helpfulness was not altered with variance due to response time removed.
Figure 3. Mean compliance level, expressed as number of hours pledged, exhibited by women (top panel) and men (bottom panel) in Experiment 2.

175) = 2.57, p < .05. (The main effect of sex was also reliable, F(1, 175) = 4.49, p < .05, with women pledging more time overall than men, M = .63 vs. .37.) The condition simple main effect was reliable for men, F(4, 94) = 5.45, p < .01, but not for women (F < 1). For men, one experimental group-control group contrast was reliable: more time was pledged in the extended-thanks-consensus condition (M = 1.00) than in the control condition (M = 0.40), t(175) = 2.50, p < .01—this amounted to the only FITD effect found in Experiment 2.

Compliance levels, we noticed, were much lower in general in Experiment 2 than in Experiment 1, a fact confirmed when the mean overall compliance level of the former (M = .87) was compared to that of the latter (M = .49), t(318) = 3.47, p < .001. More specific comparisons were also made. As noted above, the extended-thanks-no-consensus, thanks-alone, and baseline conditions of Experiment 2 involved the same procedures as, respectively, the extended-thanks, thanks-alone, and baseline conditions of Experiment 1. Mean compliance levels are given in Table 1. A 2 X 3 (experiment X condition) ANOVA was performed to test for a possible experiment X condition interaction, which turned out to be reliable, F(2, 238) = 3.30, p < .05. In comparing means of similar conditions across experiments, one contrast was reliable, and highly so. Extended thanks in Experiment 1 was associated with higher compliance than extended thanks in Experiment 2, t(238) = 4.19, p < .001.

Correlations. As in Experiment 1, we began by assessing the overall and within-cell correlations between each self-perception index (helpfulness and submissiveness ratings; helpfulness, submissiveness, and neutral latencies) and compliance level. Here again none of the coefficients reliably exceeded zero (range of —.02 to .09 with mean of .05). For instance, the overall Pearson correlation between helpfulness ratings and compliance was .09. Similar correlational analyses were carried out for each gender separately. None of these was reliable. For example, the correlation between helpfulness ratings and compliance was .14 for women and .01 for men.

We also pursued the partial correlation and composite score approaches carried out in Experiment 1 and as before found no reliable relationships between self-perception and compliance. In yet a further approach, within cell correlations between each self-perception measure (including the helpfulness composite score) and hours pledged were pooled in two ways: (a) over all eight major conditions comprising both experiments (n = 320) and (b) over the two cells in which heightened compliance was evident (i.e., extended thanks in Experiment 1 and extended-thanks-consensus for males alone in Experiment 2). None of the correlations was reliable.

With respect to the other individual-difference measures and their relationship to compliance, expectations about future experimental participation and hours pledged correlated reliably overall, r(183) = .30, p < .05, and tended to correlate highest within consensus conditions: extended-thanks-consensus, r(38) = .54, p < .01; extended-thanks-no consensus, r(36) = .05, ns; thanks-alone-consensus, r(34) = .42, p < .01; thanks-alone-no consensus, r(31) = .18, ns; and baseline, r(36) = .27, ns. We found no correlation between compliance and mood, positive r(183) = .08, ns, or negative r(183) = .06, ns.

7 The analyses performed on compliance data were also carried out with scores of participants who refused the small request eliminated. Results were not changed in any material way by this exclusion.
Discussion

Self-perception change. Our first comments concern what further we have learned about how a self-perception process is activated in the FITD situation. Helpfulness ratings were faster than baseline in the two experimental conditions in which consensus information was withheld (i.e., extended thanks-no consensus, thanks alone-no consensus). These data suggest that so long as something relevant to one's status as a helpful person occurs— in this case, the opportunity to consent to perform a small favor—the self-evaluation process is triggered. Responses will be relatively fast if all the available information defines the participant as helpful. When some information clearly contradicts this notion, as it did in the two consensus information conditions (i.e., extended-thanks-consensus and thanks-alone-consensus), time is required to resolve the inconsistency and, as a result, response times get longer.

The pattern of response times we saw in Experiment 1 was replicated in Experiment 2. Procedures used in the extended-thanks, thanks-alone, and baseline conditions of Experiment 1 were the same as those used in the extended-thanks-no-consensus, thanks-alone-no-consensus, and control conditions, respectively, of Experiment 2. Again, latencies were shorter than baseline in the extended-thanks (no consensus) and thanks-alone (no consensus) conditions, which again contradicts the version of the self-perception account of FITD to propose that extended thanks are necessary for arousing a self-perception process. The process appears to take place so long as a small request is consented to and brings at least a thank you.

Once set in motion, the self-perception process can be directed subtly by the experimenter. One means of influencing how participants rate themselves is by means of the feedback given for compliance with the small request. If the feedback un-equivocally implies that the participant has been helpful, as is the case when extended thanks is provided in the absence of consensus information, helpfulness will be rated quite high. If, by contrast, feedback implies that the participant has not been particularly helpful, as is the case when consensus information without an extended show of thanks is provided, helpfulness will be rated relatively low.

Self-rating findings of Experiment 1 were partially replicated. As in the earlier study, in Experiment 2 extended feedback (no consensus) led to higher-than-baseline levels of self-rated helpfulness. However, unlike Experiment 1, thanks alone (no consensus) failed to increase the ratings above baseline. Extended thanks reliably carried a clear message of gratitude—perhaps of more gratitude than participants expected, given the magnitude of the favor that was being asked for. Accordingly, self-ratings were adjusted upward to fit the degree of helpfulness implied by the feedback. The discrepancy between studies in the effects of thanks alone may reflect the fact that the amount of gratitude connoted by a thank you can vary according to method of delivery and circumstance. It is possible that some uncontrolled factor(s) of this sort varied systematically across our experiments to produce the differences observed in how a thank you affected self-rated helpfulness.

Summarizing to this point, participants' self-labeling was affected by observations of behavior and behavioral context. In this, self-perception theory (Bern, 1972) predicted well our latency and self-rating findings.

It is also true that experimentally induced self-perception changes were anchored to already existing self-views, such as those associated with gender. Again (as in Experiment 1) we found clear gender differences in self-perceived helpfulness, with women making higher ratings, and making them faster, than men. Our two studies demonstrate that helpfulness is a more salient aspect of the self-concepts of college women than of college men. However, we did not replicate the finding of Experiment 1 in which women rated themselves more submissive than men. Submissiveness may be an unstable aspect of the self-view due to the trait's negative connotations.

Self-perceptions and compliance. In Experiment 2, compliance with the large request occurred at an overall rate approximately half that of Experiment 1. We attribute this to our efforts in Experiment 2 to ensure the privacy of responses to the large request. The facilitating role of social pressure in compliance is well demonstrated (e.g., Milgram, 1974).

More important than the difference between the two studies in overall compliance rate is the fact that level of social pressure appears to have interacted with our FITD manipulations in influencing number of hours pledged. Under the public conditions of Experiment 1, participants who received extended thanks for their compliance with the small request complied with the large request at a higher rate than baseline. Under the private conditions of Experiment 2, by contrast, compliance with the large request in the extended-thanks condition failed to exceed baseline. It is likely that the experimenter's effusive show of gratitude created in participants an obligation to appear grateful in return. Participants concerned exclusively with impression management would alter their behavior to meet observers' expectations about the reciprocation of gratitude when an opportunity for public observation of the behavior was possible (as in Experiment 1), but not otherwise (as in Experiment 2). Consistent with this interpretation, reciprocation of extended thanks only occurred under the public evaluation conditions of Experiment 1. This further contradicts the self-perception account, which holds that a private concern with self-definitional maintenance is what drives compliance.

Overall, women pledged more time than did men in response to the large request, and, recall, women perceived themselves as more helpful than did men; both findings are consistent with Experiment 1. This stable relationship between preexperimental self-perceptions and large-request compliance stands in contrast to the failure of experimentally induced self-perception changes to carry through and affect compliance. We appear to have two sorts of self-perception, one that is associated with gender and that is related to compliance, and one that is modified by the context and that is unrelated to compliance. We will return below to a characterization of each type of self-perception and to an explanation for why one type affects compliance and the other fails to do so.

Men and women also differed in how the manipulations of Experiment 2 affected their compliance with the large request. Whereas women complied at a constant level over conditions, men complied more in the extended-thanks-consensus condition than they did in any other condition, experimental or base-
line. Therefore, the lone FITD effect to emerge in this investigation is embodied in the higher than baseline rate of compliance exhibited by men who received the feedback that contained extended thanks and consensus information. This particular FITD effect would have been considered a low-probability occurrence from the self-perception theory perspective; we interpreted it to imply that feedback containing extended thanks, but excluding consensus information, would have been the best way of enhancing compliance in Experiment 2.

In the only other research to investigate the role of consensus information in the FITD situation, DeJong (1981) found higher than baseline rates of compliance in a high consensus condition but did not find such elevation in a low-consensus condition. This effect is opposite to what the self-perception account would predict. In Experiment 2, when a FITD effect did occur, it also involved a condition in which consensus information was present (i.e., among men in the extended-thanks-consensus condition) and emerged in the absence of an associated self-perception change. But why, if not for reasons of self-perception, did this higher than baseline level of compliance occur, why only among men, and why only when extended thanks was also present? One possible explanation relies on a demonstrated gender difference in distributive justice style. In allotting rewards, men tend to use an equity norm (rewards proportional to input) whereas women tend to use equality (equal rewards for all; e.g., Major & Adams, 1983). Perhaps men in the extended-thanks-consensus condition, compared to any other condition, saw graduate students (who were the beneficiaries of large-request compliance) as highly deserving. It was in the extended-thanks-consensus condition that the experimenter (a representative of the graduate student group) showed a high level of gratitude for just a small favor, involved participants in the research process by discussing reasons why small request compliance rates were high (see content of consensus feedback), and spent the most time of any condition displaying this friendly, engaging behavior. Men, who tend to base reward allocation judgments on merit, saw reason in this situation to grant the experimenter's cohorts their wish of extra research time. In the other conditions of the experiment, where worthiness was not as well demonstrated, much smaller time pledges were made. Women, who tend to make more use than men of an equality principle, were less affected than men in their reward allocation decisions by information about merit. As such, our manipulations failed to affect women's pledges.

Finally, in Experiment 1, self-ratings of helpfulness made by participants in the extended-thanks and thanks-alone conditions failed to differ. We noted at the time that a difference may have actually been present but not detected due to a possible ceiling effect. If the difference in self-perception existed, it would account at least partially for the pattern of compliance findings in Experiment 1 and thereby lend support to the self-perception account of FITD. The findings of Experiment 2, while not necessarily discounting the possibility of a ceiling effect in the previous study, do discount the impact that such a difference, if it existed, had on compliance with the large request; self-rated helpfulness was higher in the extended-thanks condition than the thanks-alone condition of Experiment 2, but despite this, an associated compliance difference failed to emerge.

**Expectations about future participation.** In our correlational data, expectations about whether future experimental participation would be pleasant correlated significantly with number of hours volunteered. There are two possible explanations for this finding. One, germane to the theory of FITD, is that expectations are indeed instrumental in decisions about volunteering. This suggests that one way of getting FITD tactics to work is to aim them at changing expectations concerning the immediate, relatively transitory consequences (e.g., pleasantness) of participation rather than altering self-conceptions of a more general sort (e.g., the trait of helpfulness). It is not surprising that our manipulations, designed to change self-perceived helpfulness, failed to change expectations about the rewards of participation.

The other explanation, less pertinent to FITD theory, proposes that expectations were determined by time volunteered. We elicited expectation reports after the response to the large request, which leaves open the possibility that reports about enjoyment were based largely on hours pledged.

**General Discussion**

In two studies of the FITD effect, the self-perception account predicted well how our social approval and consensus level manipulations would affect self-ratings of helpfulness. However, induced self-rating changes failed to predict compliance with the large request; when self-perception changes occurred, compliance changes failed to occur, and when compliance changes occurred, they did so in the absence of preceding self-rating changes.

Our investigations present a challenge to the self-perception account of the FITD effect, especially as applied to relatively short interrequest delay situations. We do not necessarily dispute the fact that self-perceptions of helpfulness can affect compliance (cf. Berkowitz, 1987). However, our studies cast considerable doubt on the idea that a small-request situation of the typical type has the potency to change, in a consequential way, central and enduring aspects of the self such as the propensity to help. Good evidence for this notion exists in a comparison of experimentally induced changes in self-rated helpfulness with gender differences in self-rated helpfulness. The gender differences reflected the presence of attributes that existed for some time prior to the experiment, that were likely integrated with other aspects of personality (e.g., affect and conation), and therefore represented with more validity than experimentally created report differences how frequently participants had been helpful in the past, valued being helpful, and would be helpful in the future. Accordingly, the attribute differences, and not the induced ones, predicted actual helping behavior.

Altering attitudes that are as central to the self-concept as those associated with helpfulness is a tall order, not likely accomplished in the typical small-request situation used in FITD research. The observed self-perception changes induced by our manipulations were, we believe, temporary and transitory. The changes failed to reflect more permanent aspects of the self, es-
pecially attitudes toward being helpful, and were therefore not instrumental in affecting responses to the large request.8

**A Model of the FITD Compliance Enhancement Process**

If not self-perception, what does the small-request situation change that is crucial to compliance with the large request (i.e., why has the FITD effect been significant in so many studies)? In brief, the FITD situation appears to impart information about the situation (cf. Rittle, 1981), especially norms of conduct that apply and opportunities that exist for advancing personal goals. In order to explain more fully how the small-request situation functions, we propose the following model. Three changes must be brought about in the recipient before FITD will occur.

**Attitude activation.** Assume the large favor that is targeted is a donation to the cancer society. It is the source's first task to define an attitude that would drive the desired behavior; for example, the desire to be or to appear compassionate. The attitude should also be accessible in nature. An accessible attitude is one that is strong and stable, and one capable of being elicited automatically on mere presentation of the attitude object (Fazio, 1990). If, when making the large request, the source successfully arouses in the recipient the desire to show compassion, then the likelihood of a donation will be enhanced. Notice that the emphasis in this model is on eliciting powerful attitudes that already exist in the self.

**Interpretation of the large request.** Whether a targeted accessible attitude is actually called forth depends on how the large-request situation is interpreted. Some interpretations will contain the attitude object whereas others will not. Construing a request for a donation to the cancer fund as an opportunity to display compassion will have a much different effect than thinking of the request as an opportunity to lose the grocery money. The small-request situation can be crucial in determining interpretation.

**Learning.** The targeted interpretation (which is necessary for eliciting the targeted accessible attitude) can be made more probable if the experimenter disseminates certain information in the small-request situation. The information should concern norms or opportunities or both (cf. Dillard, 1990). For example, a prospective donor may be asked to accept a pamphlet about children's cancer. The lesson here is that a refusal to make a donation will be seen as especially callous because it signifies a rejection of sick children. This information would subsequently affect how a later request for a cancer donation would be defined and, consequently, determine the attitudes brought into play.

It bears mentioning that, while the small-request situation may be ill-suited to changing the self-concept, it affords the source a substantial degree of flexibility for imparting information about the situation. There are two reasons for this versatility. First, three distinct phases of the small-request situation each contribute to what the participant learns: (a) the source making a request, (b) the recipient responding (usually with compliance), and (c) the source reacting (e.g., with some level of approval or disapproval) to the recipient's response. The request, the recipient's behavior, and especially the feedback given for compliance (or refusal) can vary in numerous ways and therefore convey many different sorts of messages. Second, the source has almost total control over what happens during the first and third phases, so that the situation can be used to communicate diverse messages desired by the experimenter.

To illustrate further the three-target account, we use it to explain the FITD effects that occurred in our experiments. The first of these was the higher than baseline compliance level observed in the extended-thanks condition of Experiment 1. The small request situation in that condition consisted of the experimenter requesting a small favor, the recipient responding with consent, and the experimenter responding in turn with an extended show of gratitude. How might recipients interpret the extreme gratitude from the experimenter? The high degree of gratitude for consent to a fairly minor favor was unexpectedly responsive. Recipients might conclude from such out of proportion positive feedback that the experimenter places high value on socially responsive conduct, that the experimenter has overestimated the recipient's helpful intent in consenting to perform the small favor, or both. This information puts certain behavioral pressures on the recipient: to follow the norm of high social responsiveness, to verify the experimenter's exaggerated belief about the recipient's generosity, or both.

Against the background of what was learned in the small-request situation, a subsequent large request will be interpreted as an opportunity for reciprocation of generosity, verification of the experimenter's impressions, or both. Indeed, reciprocating generosity and keeping to a publicly perceived identity are each behaviors human beings evaluate positively and for which accessible positive attitudes tend to be possessed by most people (cf. Cialdini, 1993). Consequently, these attitudes will be activated in the large-request situation, manifest themselves in a desire to perform the favor that is requested, and lead to performance of the favor.

The other FITD effect consisted of the higher than baseline rate of compliance exhibited by men in the extended-thanks-consensus condition of Experiment 2. It was in this condition that consent to perform the small favor brought both extended thanks and consensus information from the experimenter. This feedback was the lengthiest of all types used and, more than in any other condition, depicted an attentive and concerned experimenter. Against this background, how might the large request be interpreted? Men, who tend to think in terms of equity, would evaluate the experimenter's attentive behavior as meritorious and deserving. The request to help the graduate students, a group the experimenter belonged to, was an opportunity for the men to act on their equity concerns by donating research time, which is what they did. This account also suggests that attitudes toward private objects (e.g., being equitable), apart

---

8 This is not to say that some sort of FITD strategy could not be used to create significant changes to the self. One can see how the escalating commitments induced by captors in a prisoner of war camp (e.g., Cialdini, 1993) or by the infamous Reverend Jim Jones in his jungle commune (Osherow, 1992) could produce self-perception changes that subsequently could help drive compliance. However, these phenomena occurred over many trials and involved the arousal of cognitive dissonance.
from attitudes toward public objects (e.g., appearing consistent), can affect compliance with the large request.

To conclude: Over many years of FITD research, it has been assumed that self-perception change is the crucial mechanism of compliance enhancement. However, our data suggest that changing self-perceptions and sustaining these changes may be a difficult task using FITD tactics. Although temporary changes are relatively easy to bring about, such changes have no impact on compliance. By contrast, FITD tactics aimed at changing situational definitions may be the mechanism of compliance enhancement. We expect that when the FITD situation changes interpretations that activate attitudes that enhance compliance, the prospects for producing FITD effects are better than if the small-request situation is aimed at creating central changes in the self-concept.

References


Received February 28, 1994
Revision received August 8, 1994
Accepted January 26, 1995