Five Studies Testing Two New Egoistic Alternatives to the Empathy–Altruism Hypothesis

C. Daniel Batson, Janine L. Dyck, J. Randall Brandt, Judy G. Batson, Anne L. Powell, M. Rosalie McMaster, and Cari Griffitt
University of Kansas

The empathy–altruism hypothesis claims that prosocial motivation associated with feeling empathy for a person in need is directed toward the ultimate goal of benefiting that person, not toward some subtle form of self-benefit. We explored two new egoistic alternatives to this hypothesis. The empathy-specific reward hypothesis proposes that the prosocial motivation associated with empathy is directed toward the goal of obtaining social or self-rewards (i.e., praise, honor, and pride). The empathy-specific punishment hypothesis proposes that this motivation is directed toward the goal of avoiding social or self-punishments (i.e., censure, guilt, and shame). Study 1 provided an initial test of the empathy-specific reward hypothesis. Studies 2 through 4 used three procedures to test the empathy-specific punishment hypothesis. In Study 5, a Stroop procedure was used to assess the role of reward-relevant, punishment-relevant, and victim-relevant cognitions in mediating the empathy–helping relationship. Results of these five studies did not support either the empathy-specific reward or the empathy-specific punishment hypothesis. Instead, results of each supported the empathy–altruism hypothesis. Evidence that empathic emotion evokes altruistic motivation continues to mount.

Feeling empathy—defined as an other-oriented emotional response congruent with the perceived welfare of another person—can evoke motivation to help that person (Coke, Batson, & McDavis, 1978; Eisenberg & Miller, 1987; Krebs, 1975). Several researchers have suggested that this motivation is, at least in part, altruistic; they claim that empathy evokes motivation directed toward the ultimate goal of benefiting the person for whom empathy is felt, not toward some subtle form of self-benefit (see Batson, 1987; Batson, Duncan, Ackerman, Buckley, & Birch, 1981; Hoffman, 1976; Krebs, 1975). If valid, this empathy–altruism hypothesis seems very important. It contradicts the general assumption in psychology that all motivation, including all prosocial motivation, is ultimately egoistic (Wallach & Wallach, 1983).

Initial support for the empathy–altruism hypothesis comes from a series of studies using Empathy × Ease of Escape designs. Researchers have consistently found that when empathy is low, helping drops dramatically if escape is easy. When empathy is high, however, helping remains high even if the empathically aroused individuals can easily reduce their arousal by escaping exposure to the suffering victim (Batson et al., 1981; Batson, O’Quin, Fultz, Vanderplas, & Isen, 1983; Fultz, Batson, Fortenbach, McCarthy, & Verney, 1986; Toi & Batson, 1982). This pattern of results seems to rule out the most popular egoistic explanation of the empathy–helping relationship: that the empathically aroused individual helps in order to reduce his or her own aversive empathic arousal (Hoffman, 1981; Piliavin & Piliavin, 1973).

NEW EGOISTIC ALTERNATIVES TO THE EMPATHY–ALTRUISM HYPOTHESIS

Even if the motivation to help associated with empathy is not directed toward the goal of aversive-arousal reduction, it may still be egoistic. Archer, Díaz-Loving, Gollwitzer, Davis, and Foushee (1981), Batson (1987), Cialdini et al. (1987), Dovidio (1984), and Thompson, Cowan, and Rosenhan (1980) have each proposed possible new egoistic explanations for all or part of the evidence presented in support of the empathy–altruism hypothesis. In essence, these new egoistic explanations suggest that apart from any general rewards and punishments associated with helping, there are empathy-specific rewards or punish-
ments. Feeling empathy changes the anticipated reward or punishment structure of the situation, making helping more beneficial to the self and so increasing the egoistic motivation to help.

Empathy-Specific Reward Hypothesis

According to one new egoistic explanation, people learn through prior reinforcement either that (a) special rewards in the form of praise, honor, and pride are attendant on helping when feeling empathy (Batson, 1987; Thompson et al., 1980; see also Meindl & Lerner, 1983) or (b) empathy creates a special need for these rewards of helping (Cialdini et al., 1987). When people feel empathy, they think of these social and self-rewards and are egoistically motivated to obtain them.

This empathy-specific reward hypothesis can easily account for the results of the studies cited earlier that provide support for the empathy–altruism hypothesis, because all of those studies involved manipulation of ease of escape. Empathy-specific rewards should accrue only when one helps, so the empathy-specific reward hypothesis predicts that a chance to escape will not reduce helping by empathically aroused individuals. No reduction in helping under easy escape is precisely what has been found.

Empathy-Specific Punishment Hypothesis

According to a second new egoistic explanation, feeling empathy may lead to increased helping because of anticipated empathy-specific punishments for failing to help (Archer et al., 1981; Batson, 1987; Dovidio, 1984). Presumably, people learn through prior reinforcement that a special obligation to help—and special guilt and shame for not helping—are attendant on feeling empathy. When people feel empathy, they think of these social and self-punishments and are egoistically motivated to avoid them.

This empathy-specific punishment hypothesis also can account for the results of the studies that seem to support the empathy–altruism hypothesis. The empathy-specific punishment hypothesis predicts reduced helping by empathically aroused individuals when it is easy to escape the anticipated punishments for failing to help. However, escape from the victim’s suffering, which was the form of escape made easy in the studies supporting the empathy–altruism hypothesis, might not have allowed escape from the anticipated guilt and shame for doing so. Even in the easy-escape condition of those studies, empathically aroused individuals might have found it difficult to escape these anticipated punishments.

We report five studies designed to test these two new egoistic alternatives to the empathy–altruism reward hypothesis. Study 1 provided an initial test of the empathy-specific reward hypothesis. Studies 2, 3, and 4 used three procedures to test the empathy-specific punishment hypothesis. In Study 5 we used a different procedure to test both new egoistic alternatives, as well as the empathy–altruism hypothesis.

TESTING THE EMPATHY-SPECIFIC REWARD HYPOTHESIS: EFFECTS ON MOOD OF NOT BEING ALLOWED TO HELP

Social and self-rewards associated with helping produce enhanced mood (Yinon & Landau, 1987) as long as the helping act has a low cost and is clearly beneficial to the person in need (Weyant, 1978). This fact suggests that one way to test the relative merits of the empathy–altruism and the empathy-specific reward hypotheses is to compare (a) the mood of individuals who believe that the need of a person for whom they feel empathy has been relieved as a result of their own action with (b) the mood of individuals who believe that the person’s need has been relieved, but not as a result of their own action. Because the mood-enhancing rewards for helping should be available only to the helper, the empathy-specific reward hypothesis predicts that individuals feeling a high degree of empathy will be in a more positive mood when they have been the agent of the victim’s relief than when they have not. Moreover, to the extent that empathy-specific reward is the only motivational process operating, the mood of high-empathy individuals deprived of the opportunity to help should be unaffected by whether the victim’s need is relieved by other means. Neither relief of the need nor lack of relief per se is relevant to the egoistic goal of obtaining mood-enhancing rewards for helping.

In contrast, the empathy–altruism hypothesis predicts that individuals feeling a high degree of empathy will be in as positive a mood when the victim’s need is relieved by other means as when by their own action. Moreover, the empathy–altruism hypothesis predicts that when empathically aroused individuals are deprived of the opportunity to help, they should be in a more positive mood when the victim’s need is relieved by other means than when it is not relieved. Relief, by whatever means, brings attainment of the altruistic goal of increasing the victim’s welfare. This goal attainment should increase positive mood.

Manipulation and Measurement of Empathy

In the studies providing evidence for the empathy–altruism hypothesis, empathy sometimes has been experimentally manipulated (Batson et al., 1981, Experiments 1 and 2), sometimes allowed to occur naturally and then measured through self-reports (Batson et al., 1983, Studies 1 and 2; Batson, Bolen, Cross, & Neuringer-Benefiel, 1986; Fultz et al., 1986, Study 1), and sometimes both manipulated and measured (Coke et al., 1978, Experiment 2; Fultz et al., 1986, Study 2; Toi & Batson, 1982). Exactly the same pattern of helping, the pattern predicted by the empathy–altruism hypothesis, has been found when empathy is manipulated and when it is measured, suggesting the interchangeability of these two techniques for operationalizing empathy as an independent variable.

In most cases, of course, experimental manipulation of an independent variable is preferable to measurement because manipulation permits clearer causal inference. In our case, however, measurement of empathy actually seemed preferable to manipulation. Our goal was to determine whether the empathy-specific reward hypothesis merited further consideration, so we wished to give this hypothesis the best possible chance to display its power. Because this hypothesis was based on a socialization model that implied individual differences and awareness of one’s empathic state, the predicted self-reward effects seemed most likely to appear among subjects who reported themselves to be feeling empathy.

If high-empathy subjects report a more positive mood when
they are the agent of the victim's relief than when they are not, as the empathy-specific reward hypothesis predicts, then additional research would be required to ensure that feeling empathy, and not some correlate of feeling empathy, causes the effect. If, however, these subjects fail to report a more positive mood when they are the agent of the victim's relief, this finding would count against the empathy-specific reward hypothesis as clearly as failure to find this predicted effect in an experimental design in which empathy is manipulated. As Campbell and Stanley (1963) pointed out long ago, failure to find predicted effects in a correlational design counts against a causal hypothesis as clearly as failure to find these effects in an experimental design. Given both this equal power to detect lack of support for the hypothesis and our desire to optimize chances of finding empathy-specific reward effects if they exist, it seemed best in this case to operationalize empathy through measurement, not manipulation.

Study 1: Empathy and Not Being Allowed to Help

Design and Predictions

To test the predictions of the empathy-specific reward hypothesis, we created a situation in which all subjects were (a) confronted with a person in need and (b) informed that they could perform a task that would, at little or no cost to them, relieve the victim's need. (Low-cost helping was necessary because research by Weyant, 1978, and Yinon & Landau, 1987, suggested that the mood-enhancing effects are limited to low-cost, clearly beneficial help.) Subjects then completed a measure of self-reported empathic emotional reaction to the victim, followed by an initial measure of mood. Later, half of the subjects learned that by chance they would not be performing the task that could, at little or no cost to them, relieve the victim's need. (No-cost helping was necessary because research by Weyant, 1978, and Yinon & Landau, 1987, suggested that the mood-enhancing effects are limited to low-cost, clearly beneficial help.) Subjects then completed a measure of self-reported empathic emotional reaction to the victim, followed by an initial measure of mood. Later, half of the subjects learned that by chance they would not be performing the helping task after all. Moreover, both among subjects allowed to perform the task and those not allowed, half learned that the victim was still in need, and half learned that by chance the victim was no longer in need. Finally, subjects completed a second measure of mood. A median split on the measure of self-reported empathy was combined with the two experimental manipulations to produce a 2 (low vs. high empathy) × 2 (no prior relief of victim's need vs. prior relief) × 2 (perform the helping task vs. not perform) factorial quasi-experimental design.1

The major dependent measure was change in self-reported mood after subjects were or were not allowed to help the victim. The empathy-specific reward hypothesis predicted a 1 versus 3 pattern of mood change among high-empathy subjects: In the no-prior-relief/perform cell, mood should not change or should become more positive, whereas in each of the other three cells—in which subjects had in different ways been deprived of the anticipated opportunity to obtain the empathy-specific rewards for helping—mood should become more negative. The empathy–altruism hypothesis also predicted a 1 versus 3 pattern of mood change among high-empathy subjects, but a different one: In the no-prior-relief/not-perform cell, mood should become more negative, whereas in each of the other three cells—in which in one way or another the victim's need had been relieved—not change or should become more positive.

Method

Subjects

Eighty students (40 men, 40 women) in an introductory psychology course at the University of Kansas served as subjects, partially fulfilling a course requirement. Within sex, a randomized-block procedure was used to assign 20 subjects (10 men, 10 women) to each of the four cells of the 2 (no prior relief of victim's need vs. prior relief) × 2 (perform the helping task vs. not perform) experimental design.2

Procedure

Subjects were run individually. A written introduction described the study as concerning "how a variety of task characteristics and outcome consequences affect people's task performance and attitudes." Four task characteristics were being examined: (a) for whom one's performance has consequences—self only, another person only, or both; (b) the kind of outcome consequences—positive, negative, or neutral; (c) stability of the task situation—stable or unstable in opportunity to perform the task, outcome consequences, or both; and (d) complexity of the task situation—simple (varying only in whether the consequences for self are positive, negative, or neutral) or complex (varying in type of consequences, for whom, and stability).

Ostensibly, 2 same-sex subjects were participating in the study. One would be randomly assigned to the simple task situation and the other to some variant of the complex task situation. The 2 subjects were not to meet face-to-face during the study, but the one assigned to the complex task situation was to receive an audio communication from the one assigned to the simple task situation.

The outcome consequences—positive, negative, and neutral—were then described. Positive consequences involved either receiving a raffle ticket or avoiding an electric shock for each correct response. Negative consequences involved receiving a "mild but uncomfortable" electric shock for each incorrect response. Neutral consequences involved simply being informed whether each response was correct or incorrect.

Finally, the introduction explained that unstable task situations could vary either in opportunity to perform the task or in outcome consequences. Variance in opportunity to perform was illustrated by second-string players on sports teams and understudies in the theater, "people who must keep their skills honed and be ready to take over when necessary but who may never get a chance to perform." Variance in outcome consequences was illustrated by the change from positive to negative consequences that occurs when an underdog sports team becomes the favorite. The former may have nothing to lose, whereas the latter has much to lose.

Condition 9: benefiting the other person; unstable task situation. Once they had read this information, all subjects learned that they had (ostensibly randomly) been assigned to Condition 9, a complex task situation in which they were to perform a task with positive consequences

---

1 In the prior-relief condition, the task could no longer benefit the victim, so it may seem inappropriate to label it a helping task. We have retained this label for two reasons. First, the task was originally presented to subjects as one on which their performance could help the victim. Second, we wish to emphasize that the task was the same in both the no-prior-relief and the prior-relief conditions and that only the consequences of the task were different.

2 Ten additional students (6 men, 4 women) were excluded from the sample and replaced because of suspicion. The suspicion rate did not differ reliably across experimental conditions, \( x^2(3, N = 90) = 6.08, p > .10 \). Moreover, analyses the same as those reported here but including the 10 suspicious individuals produced the same, albeit somewhat weaker, pattern of reliable results.
for another person under unstable conditions. A sheet provided more detailed information about this condition:

Your performance will have no consequences for yourself; it will have positive consequences for the other research participant. . . . That person has been initially assigned to the negative consequences condition of the simple task situation, which means that he or she will receive an electric shock after every error he or she makes (most people make about 9-10 errors in the simple task situation). But you can help that person avoid the negative consequences. For every correct response you make on your task, the person in the simple task situation will be given neutral consequences—just information—instead of shocks after one of his or her errors. If you make enough correct responses, he or she will receive no shocks at all.

Because the task situation was unstable, there was a 1 in 3 chance that either the opportunity to perform the task or the outcome consequences or both would change just before time to perform the task.

This sheet also explained that before performing their task subjects would complete a practice task “like the one you will be asked to perform” and would then “get to know” the person assigned to the simple task situation by listening as he or she talked to them briefly over an audio intercom. Finally, it was explained that the other person was not aware of the subject’s opportunity to help him (her) avoid the shocks.

Learning of the other person’s need. After performing the practice task and receiving positive feedback (to allay fears about not being able to succeed at the helping task), subjects were left alone to listen to the other subject—Brian for men, Janet for women—talk over the intercom. What they heard was actually a prerecorded tape. Explaining that he (she) was supposed to talk about “what’s on my mind right now,” Brian (Janet) hesitated and said:

Well . . . , I guess if I’m really honest, I’d have to say I’m thinking about this shock thing. I mean, when I came in here I didn’t really expect anything quite like this. . . . When I found out I’d get shocked for every mistake I make on my task, I wasn’t crazy about the idea, but I thought, it’s probably no big deal. Well . . . , they gave me a couple of sample shocks so that I’d know what to expect. . . . Wow! (nervous laugh) I don’t want to sound like a wimp or anything, but I was surprised. Those shocks kinda hurt! I mean, they weren’t terrible and I guess I’ll go through with it, but I’m not looking forward to making mistakes on my task. And I guess from what they say I’ll probably make some. . . . Oh well, I’m sure it won’t be that bad.

Empathic reaction to learning of the other’s need. After listening, subjects completed an emotional response questionnaire, which was designed to assess their empathic feelings toward Brian (Janet). This questionnaire listed 24 adjectives describing different emotional states. Subjects were to indicate on 7-point scales (1 = not at all, 7 = extremely) how they were feeling. Included in the list were 6 adjectives that had been previously found through factor analysis (in six studies; see Batson, 1987, for a review) to reflect feelings of empathy: sympathetic, moved, compassionate, tender, warm, and softhearted.

Measure of mood prior to experimental manipulations. Once subjects completed the emotional response questionnaire, they were given the first mood measure. This measure consisted of fifteen 9-point bipolar scales; subjects were to indicate their present feelings by circling the number on each scale “that best represents how you are feeling right now.” Seven of the scales were used to provide a measure of the evaluative tone of subjects’ mood: bad mood—good mood, sad—happy, depressed—elated, dissatisfied—satisfied, gloomy—cheerful, displeased—pleased, and sorrowful—joyful. The first 4 items in this list had been used by Rosenhan, Saltzer, and Hargis (1981); the last 3 were added on the basis of face validity. The other eight scales were related to either the tension (e.g., nervous—calm, tense—relaxed) or potency (e.g., lethargic—energetic, passive-active) dimensions of mood; they served as filler items. (Factor analysis of our subjects’ responses revealed that the seven evaluative scales defined a mood dimension orthogonal to the tension and potency dimensions.)

Experimental manipulation of prior relief of need and of opportunity to perform. While subjects completed this first mood measure, the experimenter went to “consult the random number table to see whether there are any changes in either your opportunity to perform the task or your outcome consequences.” The experimenter, who was blind to subjects’ empathy scores throughout the procedure, had up to this point also been blind to the experimental condition. The experimenter now checked the subject’s condition, selected a prepared sheet describing the task characteristics for that condition, returned, and asked the subject to read the sheet carefully.

Prior relief of the victim’s need was manipulated by what subjects read on the sheet about Brian’s (Janet’s) consequences. Subjects in the no-prior-relief condition read that there was no change in the consequences: “The person in the simple task situation remains in the negative consequences condition. He or she will receive electric shock after every error.” Subjects in the prior-relief condition read that Brian’s (Janet’s) consequences had changed: “The person in the simple task situation has been reassigned to the neutral consequences condition. He or she will receive no shocks but will simply be informed whether a response is correct or incorrect.”

Performance of the helping task was manipulated by what subjects read about their own performance. Subjects in the perform condition read that there was no change: “You will be performing your task.” Subjects in the not-perform condition read, “You will not be performing the task.”

The two manipulations were varied factorially, producing a 2 (no prior relief of victim’s need vs. prior relief) X 2 (perform the helping task vs. not perform) experimental design. Subjects in the no-prior-relief/perform condition were the only ones who could still help the victim. For subjects in the no-prior-relief/not-perform condition, Brian (Janet) would still receive shocks for errors, but they could no longer do anything to prevent it. For subjects in the two prior-relief conditions, Brian’s (Janet’s) need had been relieved without their action.

Measure of mood after the experimental manipulations. Subjects who were still to perform the task did so. The task involved working for 120 s from left to right, line by line through a sheet filled with numbers, circling as many combinations of 13 and 47 as possible. The task was designed so that subjects could get the 9 to 10 combinations necessary to eliminate all of Brian’s (Janet’s) shocks. (Mean number of combinations circled was 11.38, SD = 1.67.) After the task, subjects in the perform condition completed a second mood measure that was identical to the first. Subjects in the not-perform condition completed this second mood measure immediately after reading the information about the change(s) in their task characteristics. As before, instructions on the mood measure emphasized that subjects were to circle the number on each bipolar scale “that best represents how you are feeling right now.”

Debriefing. Subjects were carefully debriefed, thanked for their participation, and excused.

Results and Discussion

Effect on Mood of Not Being Allowed to Help

To control for the individual differences typically found on mood measures, we assessed change of mood from the point at which all subjects believed that they would have a chance to help to the point at which some had and some had not been allowed to help. First, ratings on the seven evaluative-tone mood scales that subjects completed prior to the introduction of the experimental manipulations were averaged to form an index of mood at the point that all subjects believed that they would have
a chance to help (Cronbach's alpha = .90). In all conditions, the average mood at this point was moderately positive, overall $M = 6.01$ on the 9-point scale (1 = bad mood, 9 = good mood), with no reliable differences among the four experimental conditions (all $F_s < 1.0$).

Next, ratings on the same seven scales after the introduction of the experimental manipulations were averaged to form an index of mood after subjects either had or had not been allowed to help (Cronbach's alpha = .94). At this point too, the average mood was moderately positive ($M = 6.35$), although across the four experimental conditions there was a significant Prior Relief $\times$ Perform interaction, $F(1, 76) = 4.25$, $p < .05$, suggesting that the experimental manipulations had affected subjects' moods (both main effect $F < 1.20$). The mean response on this postmanipulation mood index for subjects in each of the four experimental conditions is reported in Table 1.

**Empathic Response to the Other Person's Need**

To determine whether the mood effects of the experimental manipulations reflected in Table 1 were those predicted by either the empathy-specific reward or the empathy–altruism hypothesis, we had to take into account the level of empathy reported in response to Brian's (Janet's) need. Therefore, we constructed an empathy index by averaging each subject's ratings of the six adjectives on the emotional response questionnaire that have been found in previous research to reflect empathy: sympathetic, moved, compassionate, tender, warm, and soft-hearted (Cronbach's alpha = .83). Scores on this empathy index were moderately high ($M = 3.81$) and variable ($SD = 1.27$) on the 7-point response scale ($1 = $not at all, $7 = $extremely). We then performed a median split on the empathy index ($Median = 3.90$). This split produced the 2 (low vs. high empathy) $\times$ 2 (no prior relief vs. prior relief) $\times$ 2 (perform helping task vs. not perform) design needed to test the predictions of the empathy-specific reward and the empathy–altruism hypotheses.

**Mood Change for Low- and High-Empathy Subjects Who Were and Were Not Allowed to Help**

To assess mood change in the $2 \times 2 \times 2$ design, we created a mood change score by subtracting each subject's premanipulation mood score from his or her postmanipulation mood score.$^3$ The mean mood change for subjects in each cell of this design is reported in Table 2. The empathy-specific reward hypothesis predicted different mood change for high- and low-empathy subjects. The empathy–altruism hypothesis predicted that mood change among high-empathy subjects would be more positive in the one cell in which they were able to gain the empathy-specific rewards for helping, the no-prior-relief/perform cell, than in the other three cells. A planned comparison contrasting the mood change in this cell with the change in the other three high-empathy cells did not support this prediction, $F(1, 72) < 1.0$. Pairwise comparisons revealed that, as predicted, the mood change was more positive in the no-prior-relief/perform cell than in the no-prior-relief/not-perform cell, $t(72) = 1.70$, $p < .05$, one-tailed. Contrary to the predictions of the empathy-specific reward hypothesis, however, there was no reliable evidence of a more positive mood change in the no-prior-relief/perform cell than in the other three high-empathy cells.

### Table 1

**Mean Ratings on the Mood Index (After Introduction of the Experimental Manipulations) by Subjects Who Were or Were Not Allowed to Help: Study I**

<table>
<thead>
<tr>
<th>Status of other's need</th>
<th>Perform helping task</th>
<th>Not perform helping task</th>
</tr>
</thead>
<tbody>
<tr>
<td>No prior relief</td>
<td>6.56&lt;sub&gt;b&lt;/sub&gt;</td>
<td>5.84&lt;sub&gt;b&lt;/sub&gt;</td>
</tr>
<tr>
<td>Prior relief</td>
<td>6.29&lt;sub&gt;b&lt;/sub&gt;</td>
<td>6.73&lt;sub&gt;b&lt;/sub&gt;</td>
</tr>
</tbody>
</table>

*Note. $N = 20$ in each cell. Only subjects in the no-prior-relief/perform cell could actually help. Ratings on the seven-item mood index were on a 9-point scale (1 = bad mood, 9 = good mood). Cell means not sharing a common subscript differ, $p < .05$, by a t test.*

### Table 2

**Mean Mood Change for Low- and High-Empathy Subjects Who Were or Were Not Allowed to Help: Study I**

<table>
<thead>
<tr>
<th></th>
<th>Low empathy</th>
<th>High empathy</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Perform helping task</td>
<td>Not perform helping task</td>
</tr>
<tr>
<td>No prior relief</td>
<td>.13</td>
<td>.10</td>
</tr>
<tr>
<td>$n$</td>
<td>12</td>
<td>11</td>
</tr>
<tr>
<td>Prior relief</td>
<td>.27</td>
<td>.43</td>
</tr>
<tr>
<td>$n$</td>
<td>10</td>
<td>10</td>
</tr>
</tbody>
</table>

*Note. Positive mood-change scores indicate more positive mood after introduction of the experimental manipulations; negative scores indicate less positive mood after introduction of the experimental manipulations.*
perform cell than in either the prior-relief/perform cell, \( t(72) = 0.15 \), or the prior-relief/not-perform cell, \( t(72) = -1.82 \).

The empathy-altruism hypothesis predicted that mood change among high-empathy subjects would be more positive in the three cells in which Brian's (Janet's) need was relieved than in the one cell in which it was not (no prior relief/not perform). A planned comparison contrasting the mood change in this cell with the change in the other three high-empathy cells provided clear support for this prediction, \( F(1, 72) = 7.09, p < .02 \). Moreover, this effect appeared to be empathy specific; the same comparison among low-empathy subjects was not statistically significant, \( F(1, 72) = 0.20 \). Indeed, the 1 versus 3 comparison among high-empathy subjects predicted by the empathy-altruism hypothesis accounted for all reliable between-cell variance in mood change in the entire 2 \( \times \) 2 \( \times \) 2 design, residual \( F(6, 72) = 1.17, ns \), even though the 1 versus 3 pattern was only marginally stronger among high-empathy subjects than among low-empathy subjects, \( F(1, 72) = 2.45, .10 < p < .15 \). Pairwise comparisons revealed that the predicted differences among the high-empathy subjects were reliable between the no-prior-relief/not-perform cell and both the no-prior-relief/perform cell, \( t(72) = 1.70, p < .05 \), one-tailed, and the prior-relief/not-perform cell, \( t(72) = 3.52, p < .001 \), one-tailed. (The particularly high mood-change score in the prior-relief/not-perform cell probably reflected a combination of motives: pleasure that the victim's need was relieved and pleasure at relief from any lingering performance apprehension.) The difference between the no-prior-relief/not-perform cell and the prior-relief/perform cell was in the predicted direction but was not reliable, \( t(72) = 1.31 \).\(^4\)

### Task Performance

Further evidence that low- and high-empathy subjects differed in their goals is provided by the differences in task performance in the perform condition. Performance was assessed by the number of combinations that subjects correctly circled during the 2 min they worked on the final task. Mean performance scores for subjects in each cell are reported in Table 3. As can be seen, high-empathy subjects circled more combinations when Brian's (Janet's) welfare still depended on their performance (\( M = 12.38 \)) than when it did not (\( M = 10.20 \)), \( t(36) = 3.37, p < .001 \). In contrast, low-empathy subjects circled more combinations when the victim's welfare did not depend on their performance (\( M = 12.40 \)) than when it did (\( M = 10.83 \)), \( t(36) = 2.43, p < .02 \). An analysis of variance (ANOVA) revealed that although neither main effect was significant, both \( F(1, 36) < 1.0 \), the Empathy \( \times \) Prior-Relief interaction was highly significant, \( F(1, 36) = 16.78, p < .001 \). This pattern of results suggests that high-empathy subjects were more motivated to do well on the task when Brian's (Janet's) welfare depended on their performance, whereas for low-empathy subjects the opposite was true. The higher performance of the low-empathy subjects in the prior-relief condition, although not predicted by either of the hypotheses tested, seemed readily interpretable as a product of (a) these subjects being more self-focused and (b) prior-relief information leading them to focus more on personal performance standards (Wicklund, 1975).

<table>
<thead>
<tr>
<th>Other's need</th>
<th>Low empathy</th>
<th>High empathy</th>
</tr>
</thead>
<tbody>
<tr>
<td>No prior relief</td>
<td>( M = 10.83 )</td>
<td>( M = 12.38 )</td>
</tr>
<tr>
<td>( SD = 1.27 )</td>
<td>( SD = 1.51 )</td>
<td></td>
</tr>
<tr>
<td>( n = 12 )</td>
<td>( n = 8 )</td>
<td></td>
</tr>
<tr>
<td>Prior relief</td>
<td>( M = 12.40 )</td>
<td>( M = 10.20 )</td>
</tr>
<tr>
<td>( SD = 0.97 )</td>
<td>( SD = 1.87 )</td>
<td></td>
</tr>
<tr>
<td>( n = 10 )</td>
<td>( n = 10 )</td>
<td></td>
</tr>
</tbody>
</table>

### Implications of Study 1

The empathy-specific reward hypothesis predicted that high-empathy subjects would feel better if they were the cause of relief of the victim's need than if they were not: Only if they were the cause would they be in line for the mood-enhancing rewards attendant on helping. Contrary to this prediction, subjects' self-reported mood provided no evidence that high-empathy subjects felt better when the victim's need was relieved by their own action than when it was relieved by other means. Thus, our data did not support the empathy-specific reward hypothesis.

Instead, both the mood change data and the performance data were entirely consistent with the claim of the empathy-altruism hypothesis that high-empathy subjects were directed toward the goal of having the victim's need reduced. If they could obtain this goal by their own action, these subjects worked relatively hard to do so and felt relatively good about the result. Even if they were not the cause, they felt relatively good if the victim's need was relieved.

### Self-Presentation

Although our data patterned very much as predicted by the empathy-altruism hypothesis, it would be erroneous for us to conclude that they prove the validity of this hypothesis. Self-presentation (Jones & Pittman, 1982) provides a possible alternative explanation for at least some of our results consistent with the empathy-altruism hypothesis. Especially when subjects were informed that they would not be performing the helping task, their subsequent mood reports might have been affected by what they thought was the appropriate response. Subjects informed that the victim would still receive shocks might have thought they should report feeling bad; subjects informed that the victim would not receive shocks might have thought they should report feeling good. Most of the data were consistent with the predictions of the empathy-altruism hypothesis, except that mood change was not as positive among high-empathy subjects as predicted. This pattern of results might be explained by the fact that subjects in the prior-relief condition were more likely to report feeling bad if the victim would still receive shocks.

\(^4\) There was a sex difference on the empathy index; women reported more empathy for Janet (\( M = 4.24 \)) than men reported for Brian (\( M = 3.32 \)), \( F(1, 72) = 11.55, p < .001 \). Such a sex difference has often been found for self-report measures of empathy, but not for physiological measures, suggesting that the sexes differ more in the appropriateness of reporting empathic feelings than in the experience of them (Eisenberg & Miller, 1987). In this study, supplemental analyses revealed that the sex difference could not account for the observed pattern of mood change. The same pattern of significant results was found when the median split on empathy was performed separately within sex.
thought they should report feeling good. That this difference was reliable only among high-empathy subjects may be due to high-empathy subjects having a greater concern for positive self-presentation than low-empathy subjects. After all, to report that one feels a high degree of empathy for someone in need, as the high-empathy subjects had done, could itself be a product of a concern for positive self-presentation.

If self-presentation is to account for all of the support we found for the empathy–altruism hypothesis, then it must be a pervasive form of self-presentation. It must have affected not only subjects' self-reports of their mood but also their task performance. Still, the possibility that our results consistent with the empathy–altruism hypothesis were a product of self-presentation cannot be ruled out entirely.

Could self-presentation also account for our failure to find the pattern of results predicted by the empathy-specific reward hypothesis? We think not. As noted earlier, it seems possible that the mood change reported by subjects who were not to perform the task could be a product of self-presentation. However, the empathy-specific reward hypothesis also predicted a difference in mood between the no-prior-relief and the prior-relief conditions for subjects who were to perform the task. It seems implausible that the lack of difference in mood change between these two conditions was a product of self-presentation. In each of these conditions the victim's need had been relieved, so any concern to present oneself as feeling good because of this should have been the same in the two conditions. In the no-prior-relief condition, however, the relief had been caused by the subject's help; in the prior-relief condition it had not. To the degree that high-empathy subjects feel better only as a result of themselves having helped the person for whom they feel empathy, as the empathy-specific reward hypothesis claims, then we would expect our high-empathy subjects to report significantly more positive mood change in the no-prior-relief/perform cell than in the prior-relief/perform cell. They did not, t(72) = 0.15.

Thus, the lack of support for the empathy-specific reward hypothesis does not appear to be attributable to self-presentation. If, as intended, our strategy of measuring naturally occurring empathy provided maximal opportunity for any self-reward effects associated with empathy to appear, then our results seem to suggest that the motivation to help evoked by empathy is not directed toward obtaining special rewards available to the empathically aroused helper.

A Second Version of the Empathy-Specific Reward Hypothesis: Negative-State Relief

There is another version of the empathy-specific reward hypothesis, recently proposed by Cialdini et al. (1987). Rather than suggesting that special mood-enhancing rewards are associated with empathically induced helping, Cialdini et al. suggested that because feeling empathy for a person in need is a negative affective state, it produces an increased need for some mood-enhancing experience, including (but not limited to) the self-rewards following helping. In this version of the empathy-specific reward hypothesis, it is not the self-rewards that are empathy specific but the need for these rewards.

Study 1 was not designed to test the Cialdini et al. (1987) negative-state relief version of the empathy-specific reward hypothesis, and, frankly, we do not know whether our results contradict negative-state relief predictions. As far as we know, the negative-state relief literature has never addressed the question of whether, once a negative state is induced, removal of the inducing conditions is sufficient to relieve the negative state. Clearly, termination of these conditions has not been assumed to be necessary: Money, praise, or other forms of self-reward that do not remove the negative-state-inducing conditions have been assumed to work (Cialdini et al., 1987). Whether termination of the inducing conditions is sufficient is unclear.

If we assume that termination of the inducing conditions is sufficient to relieve a negative state, then the negative-state relief perspective predicts the same pattern of mood change in Study 1 as the empathy–altruism hypothesis, the pattern we observed. Therefore, a different research paradigm will be needed to provide a clear test of the relative merits of the negative-state relief version of the empathy-specific reward hypothesis and the empathy–altruism hypothesis.

TESTING THE EMPATHY-SPECIFIC PUNISHMENT HYPOTHESIS: PROVIDING JUSTIFICATION FOR NOT HELPING

The empathy-specific punishment hypothesis claims that people have learned through prior reinforcement that a special obligation to help—and special guilt and shame for failure to help—are attendant on feeling empathy. As a result, when people feel empathy they are faced with impending empathy-specific social or self-censure that is above and beyond any general punishment associated with not helping, and they are egoistically motivated to avoid these empathy-specific punishments.

Socially Administered Empathy-Specific Punishments

One version of the empathy-specific punishment hypothesis, proposed by Archer et al. (1981; see also Archer, 1984), assumes that empathy-specific punishments are socially administered. According to this version, empathy leads to increased helping only when the empathic individual anticipates negative social evaluation for failing to act in a manner consistent with his or her reported feelings of concern. To test this suggestion, Fultz et al. (1986) both manipulated and measured empathy for a person in need; they then assessed the rate of helping this person under conditions of high and low social evaluation. In the high social-evaluation condition, both the experimenter and the person in need would know if the research participant decided not to help; in the low social-evaluation condition, no one but the participant would know, not the experimenter, not even the person in need. Fultz et al. (1986) found that anticipated low social evaluation did not produce less helping than high social evaluation for either low- or high-empathy subjects. Knowing that there was no cause to worry about socially mediated punishments for a failure to help did not diminish the empathy–helping relationship.

Self-Administered Empathy-Specific Punishments

The Fultz et al. (1986) results clearly count against a version of the empathy-specific punishment hypothesis that assumes
that the relevant punishments are socially administered. It does not, however, rule out the possibility that high-empathy individuals are motivated to help to avoid self-administered punishments and negative self-evaluation (Dovidio, 1984; Hoffman, 1976; Schwartz & Howard, 1981). To do this, one must manipulate expectations of self-punishment for not helping.

How is one to manipulate expectations of self-punishment? If these expectations have been internalized to the degree that they are automatic and invariant across all situations, then manipulation seems impossible. Yet, we suspect that few people, if any, have internalized procedures for self-punishment to such a degree. Even those who reflexively slap themselves with guilt and self-reproach whenever they do wrong are likely to be sensitive to situational cues in determining when they have done wrong (see Milgram, 1963, 1974). Also, given the discomfort produced by guilt and self-reproach, we suspect that most people will not reflexively self-punish but will, if possible, overlook their failures to do good. They will dole out self-punishments only in situations in which such failures are salient and inscrutable.

If it is true that self-punishment will be avoided when possible, then expectation of such punishment after a failure to help may be effectively manipulated simply by varying ease of physical escape (i.e., the expectation of continued visual exposure to the suffering victim). Individuals who expect continued exposure should find their failure to help inescapable, so they should anticipate self-punishment. Those who expect no continued exposure should be able to put their failure out of sight and, following the old adage, out of mind, allowing self-punishment to be avoided. Batson et al. (1986) found that varying the ease of physical escape did indeed seem to have this effect on anticipated self-punishment.

If physical escape reduces anticipated self-punishment for not helping, then the studies that have used an ease of physical escape manipulation to test the empathy-altruism hypothesis against an aversive-arousal reduction explanation (Batson et al., 1981; Batson et al., 1983; Toi & Batson, 1982) may also provide a test of the empathy-specific punishment hypothesis. To the degree that these studies do provide a test, they do not offer support. These studies have consistently shown that individuals feeling high empathy for a person in need do not help less when physical escape is easy than when it is difficult.

Providing Justification for Not Helping

Although suggestive, the previous research using an ease of physical escape manipulation was not explicitly designed to test the empathy-specific punishment hypothesis. We sought to provide a more explicit test by reducing the expectation of self-punishment in a different way. If, as proposed earlier, there is leeway in interpreting a given failure to help as wrong and hence deserving of self-punishment, then the expectation of self-punishment may be reduced by providing some individuals with information that would justify not helping. We did not believe that this information could be provided directly by, for example, telling individuals not to feel guilty about not helping, because calling direct attention to the failure in this way may have the reverse effect; it may highlight the associated punishments. Instead, we wished to supply justifying information in a more subtle, indirect way. We were able to think of three ways this could be done: justifying not helping through others' inaction, justifying not helping through attributional ambiguity, and justifying not helping because qualifying to help is difficult.

Justification for Not Helping Through the Inaction of Others

One way to justify not helping, suggested by the study of social influence and social norms (Moscovici, 1985; Sherif, 1936), was to provide individuals confronted with a request for help with information about how their peers had responded to this request. If most peers had decided to help, then the belief that they too ought to help should be strengthened, leading them to anticipate more self-censure if they did not. Assuming that avoidance of this censure is their goal, helping should increase. If, however, most peers had decided not to help, then the belief that they too ought to help should be weakened, leading them to anticipate less self-censure if they did not. Assuming that avoidance of this censure is their goal, helping should decrease.

Not to be confused with diffusion of responsibility, which occurs when one knows there are other potential helpers but not whether anyone else has helped (Darley & Latané, 1968), learning that others had decided not to help should produce exclusion from responsibility. It should reduce the sense that one ought to help in the situation.

Justification for Not Helping Through Attributional Ambiguity

A second way to justify not helping, suggested by the attributional-ambiguity technique developed by Snyder, Kleck, Strenta, and Mentzer (1979), was to confront individuals with a choice between one activity that would benefit the self and another that would help a person in need. For some potential helpers, helping-irrelevant attributes of the two activities could then be highlighted, as well as information that these helping-irrelevant attributes justified choosing the activity benefiting the self. This information should make these individuals less likely to anticipate self-punishment were they to choose this activity; they could justify their choice as being due to the helping-irrelevant attributes. If avoidance of punishment is their goal, helping should decrease. Snyder et al. (1979) had found a similar attributional-ambiguity technique effective in allowing individuals to justify an act that they may otherwise consider morally wrong: avoiding a handicapped person.

Justification for Not Helping Because Qualifying to Help is Difficult

A third way to justify not helping was to make qualifying to help difficult. Imagine a person who feels sorry for someone needing a bone-marrow transplant. Imagine further that, as the empathy-specific punishment hypothesis predicts, this person anticipates feeling guilty about not volunteering to undergo the rather painful marrow-donation operation that would help the needy individual. This person should be relieved, and content, to learn of a blood-type mismatch that disqualifies him or her.
as a possible donor. Even if no other source of help is available, he or she cannot be blamed for not helping.

What if, as the empathy–altruism hypothesis predicts, this person is concerned not about avoiding guilt but about reducing the needy individual’s suffering? Learning about the blood-type mismatch should not cause relief and contentment. If no other source of help is available, the altruistically motivated individual should be upset over not being eligible.

In this example, qualifying to help is entirely out of the potential helper’s control; blood types either match or they do not. Now imagine a slightly different situation in which, rather than potential helpers simply learning that they are or are not eligible, they must perform a task requiring effort in order to qualify to help. In such a situation, how hard the potential helper tries on the qualifying task should give us a behavioral measure of whether he or she is motivated to reduce the needy individual’s suffering (which requires qualifying) or to avoid self-punishment (which does not). This should be true, however, only if failure to qualify can be justified. Failure should be justifiable if the performance standard on the qualifying task is so difficult that most people fail.

Having no basis for choosing between these three ways of providing justification for not helping, we conducted three studies, one using each technique. We reasoned that if the three studies produced consistent results, providing conceptual replication, then our confidence in the results of each would be increased.

To add generality, we used different need situations and helping responses in the three studies. We also used different techniques for operationalizing empathy. In Study 2, we manipulated empathic feelings for the person in need; in Studies 3 and 4, we measured these feelings through self-reports.

Study 2: Justification for Not Helping Through the Inaction of Previous Potential Helpers

To test the empathy-specific punishment hypothesis by providing justification for a failure to help through the inaction of previous potential helpers, we needed a research paradigm in which (a) the empathy–helping relationship was known to occur and (b) previous helping (or lack of helping) by others would not eliminate the victim’s need for help from the subject. The paradigm developed by Coke et al. (1978, Experiment 1) seemed to meet these requirements. In this paradigm, subjects learn of a young woman’s need by listening to a (bogus) pilot radio newscast and are given an unexpected chance to help her. Empathy is manipulated by instructing subjects to take a particular perspective while listening to the newscast (Stotland, 1969).

To manipulate justification for not helping, we modified the pledge form used by Coke et al. (1978) to include spaces for the responses of 8 individuals. Seven spaces were already filled, ostensibly by previous subjects. In the low-justification condition, 5 of the 7 previous subjects had volunteered to help. In the high-justification condition, only 2 of the 7 had volunteered. To ensure that we had reproduced the empathy–helping relationship, we also included two replication cells, one low empathy and one high, in which the pledge form had space only for the subject’s response. Individuals in these cells received no information about the helping of previous subjects.

Predictions

In the replication cells, both the empathy-specific punishment hypothesis and the empathy–altruism hypothesis predicted more overall helping in the high-empathy condition than in the low, because both assumed that increased empathy leads to increased helping. These two hypotheses differed, however, in their predictions for the effects of the justification manipulation on helping. The empathy-specific punishment hypothesis predicted less helping in the high-justification condition than in the low for subjects in both empathy conditions. In the low-empathy condition, subjects should be motivated to avoid general shame and guilt associated with a failure to help; in the high-empathy condition, subjects should be even more highly motivated to avoid shame and guilt because of the added empathy-specific punishments. In both conditions, it should be easier to avoid shame and guilt without having to help in the high-justification condition than in the low.

The empathy–altruism hypothesis predicted less helping in the high-justification condition than in the low for subjects in the low-empathy condition, but it predicted little or no effect of the justification manipulation in the high-empathy condition. It predicted that subjects in the high-empathy condition would be, at least partly, motivated to reduce the need of the person for whom empathy was felt; reduction in anticipated general and empathy-specific self-punishment provided by the justification information would not be relevant to reaching this goal. Across the four cells of the 2 (low vs. high empathy) X 2 (low vs. high justification) design, then, the pattern of results most consistent with the empathy–altruism hypothesis would be a 1 versus 3 pattern; the rate of helping in the low-empathy/high-justification cell should be lower than the rate in the other three cells.

Method

Subjects

For Study 2, 120 students (60 men, 60 women) in an introductory psychology course at the University of Kansas served as subjects, partially fulfilling a course requirement. By use of a randomized-block procedure, 20 subjects (10 men, 10 women) were assigned to each cell of the 2 (low vs. high empathy) X 2 (low vs. high justification) factorial design. In addition, 20 subjects (10 men, 10 women) were assigned to both the low- and high-empathy replication cells.5

Procedure

Ostensibly as part of an ongoing project for pilot testing new programs for the local university radio station, subjects listened to two pilot tapes, one for “Bulletin Board,” a program announcing campus activities, and one for “News From the Personal Side,” a program attempting to go beyond the facts of local news events “to report how these events affect the lives of the individuals involved.” Subjects were asked to adopt

---

5 Twelve additional students (8 men, 4 women) were excluded from the sample and replaced because of suspicion. Degree of suspicion did not differ reliably across the six conditions, F(5, 126) = 1.43, p > .20, and analyses the same as those reported here but including suspicious individuals produced the same, albeit somewhat weaker, pattern of reliable results.
NEW EGOISTIC ALTERNATIVES TO THE EMPATHY-ALTRUISM HYPOTHESIS

a particular listening perspective for each broadcast, and reactions were measured by questionnaires. The specific procedure was similar to that used by Coke et al. (1978, Experiment 1), except for insertion of the justification manipulation and omission of their misattribution manipulation. Therefore, we present in detail only the experimental manipulations and dependent measures.

Manipulation of empathy. Before listening to the "News From the Personal Side" tape, which presented a young woman in need, subjects were given one of two listening perspectives. Subjects in the low-empathy condition were instructed to do the following:

Try to focus on the technical aspects of the broadcast. Try to concentrate on those techniques and devices that are used to make the broadcast have an impact on the listener.

Subjects in the high-empathy condition were instructed to do the following:

Try to imagine how the person who is being interviewed feels about what has happened and how the events have affected her life. Try to feel the full impact of what this person has been through and how she feels as a result.

Subjects were led to believe that all previous research participants had been given the same listening-perspective instructions they received. This ensured that, when interpreting the justification information, subjects would perceive the other research participants to be responding to the same helping situation as they were. The experimenter was blind to which listening-perspective instructions subjects received.

Listening to the "News From the Personal Side" tape. On the "News From the Personal Side" tape, a male announcer interviewed Katie Banks, a senior at the university. Katie's parents and a sister had recently been killed in an automobile crash. Her parents did not have life insurance, and Katie was desperately struggling to support her surviving younger brother and sister while she finished her last year of college. If she did not finish, she would have to put the children up for adoption.

After the tape, the experimenter returned and prepared to administer a reaction questionnaire, only to find that an apparent mimeograph malfunction had left a blank streak across the page, rendering several response options illegible. While the experimenter went to get another copy, the subject was directed to read two letters that the professor in charge of the study had asked to be given to participants.

Katie's request for help. The first letter was from the professor. He thanked subjects for participating in the research and explained that it had occurred to him that some participants may want to help Katie, so he had encouraged her to write a letter telling what they could do if they wished. The second letter was from Katie; she presented subjects with a range of possible ways to help: sitting with her younger brother and sister while she attended her night classes, fixing things around the house, providing transportation, helping with telephone calls, and stuffing envelopes for a fundraising project.

Manipulation of justification for not helping. Enclosed with the two letters was a response form on which subjects were to indicate whether they wished to help Katie. In both justification conditions, this form had spaces for the responses of 8 subjects; the first 7 were already filled in. In the high-justification condition, only 2 of the 7 had helped: 1 subject volunteered 1 to 2 hr, the other 3 to 5 hr. In the low-justification condition, only 3 of the 7 helped: 1 subject volunteered 1 to 2 hr, the other 3 to 5 hr. (Subjects in the justification conditions filled in the last blank on the form so that they could be assured that the form with their name on it would not be seen by other subjects. This avoided concern over praise or censure from friends or acquaintances who could learn of their response.) In the replication condition, the form provided exactly the same response options as in the justification conditions but was designed for only one person's response; subjects were not provided with information about the action of their peers.

To ensure that subjects did not feel that they were Katie's last chance for help, the experimenter made it clear that theirs was not the only response form and that Katie's request would be presented to other research participants. The experimenter was blind to which version of the response form subjects received.

Dependent measure: volunteering to help Katie. Each version of the response form asked subjects to provide their name, to indicate whether they wished to help Katie and, if so, to check the number of hours they wished to volunteer: 1 to 2, 3 to 5, 6 to 8, or 9 to 10 hr. The amount of time, if any, that subjects volunteered to help Katie was the dependent measure of helping.

Ancillary measures. After subjects filled out the response form, the experimenter returned, collected the form, and gave them a legible copy of the reaction questionnaire. In addition to asking subjects how interesting, informative, and worthwhile they felt the broadcast was, this questionnaire included an item concerning Katie's need (i.e., "How great is the need of the person who was interviewed?") and two items designed to check the effectiveness of the empathy manipulation (i.e., "While listening to this broadcast, to what extent did you concentrate on the technical aspects of the broadcast?"). "To what extent did you concentrate on the feelings of the person being interviewed?").

The experimenter also gave subjects a help-opportunity questionnaire. The experimenter explained that although this questionnaire was not part of the pilot-testing study, the professor in charge wished to learn how research participants felt about being given the chance to help Katie. Included on this questionnaire was an item designed to check the effect of the justification manipulation on perceived obligation to help Katie: "Do you believe that KU students ought to help Katie?"

Debriefing. Subjects were left alone to complete these questionnaires; then they were carefully debriefed, thanked for their participation, and excused.

Results and Discussion

Perception of Katie's Need

On the reaction questionnaire, subjects indicated the magnitude of Katie's need (1 = very little, 9 = very great). Subjects in all six conditions perceived her need to be great (cell Ms ranged from 7.65 to 8.45; overall M = 8.17), with no reliable differences for either main effect or for the interaction. Apparently, the perspective-taking instructions used to manipulate empathy did not prevent subjects in the low-empathy condition from perceiving Katie to be in considerable need. Moreover, as intended, subjects perceived her need to be as great when 5 of 7 previous subjects had helped (M = 8.38) as when only 2 of 7 had helped (M = 8.05) or when there was no information about the helping of previous subjects (M = 8.08).

Effectiveness of the Empathy Manipulation

Subjects also indicated on the reaction questionnaire the extent to which they concentrated on (a) the technical aspects of the broadcast and (b) the feelings of the person being interviewed (1 = not at all, 9 = very much for each question). Across all six conditions, subjects in the low-empathy condition reported more concentration on the technical aspects (M = 6.70) than did subjects in the high-empathy condition (M = 4.55), F(1, 114) = 30.71, p < .001. In addition, subjects in the low-empathy condition reported less concentration on feelings
(M = 6.63) than did subjects in the high-empathy condition (M = 7.95), F(1, 114) = 19.31, p < .001. For neither measure was there a significant main effect for the justification manipulation or interaction, all Fs(2, 114) < 1.80. We concluded that the empathy manipulation was successful.6

**Effectiveness of the Justification for Not Helping Manipulation**

After about one third of the subjects were run, we began asking subjects during debriefing (before the true purpose of the research was revealed) for the total number of previous subjects whose names appeared on the response form and the number who volunteered to help Katie. These reports were available for 54 of the 80 subjects in the two justification conditions. Creating a proportion for each subject by dividing the reported number of previous subjects volunteering by the total and then averaging these proportions, we found that subjects in the low-justification condition recalled previous subjects helping more often (M = .72) than did subjects in the high-justification condition (M = .33), F(1, 50) = 231.53, p < .0001. These mean proportions closely approximated the proportions that subjects in the two conditions had actually been given: .71 (5 of 7) in the low-justification condition and .29 (2 of 7) in the high. Neither the empathy main effect nor the interaction was significant, F(1, 50) = .04. Neither the empathy main effect, F(1, 114) < 1.80. We concluded that the justification manipulation was successful.

**Effect of Justification on Helping**

Helping responses were coded in two ways. First, the proportion of subjects who volunteered any amount of time served as a dichotomous measure of helping (0 = no help, 1 = help); second, scores on the 5-point scale of number of hours volunteered (0 = 0 hr, 1 = 1 to 2 hr, 2 = 3 to 5 hr, 3 = 6 to 8 hr, and 4 = 9 to 10 hr) served as a continuous measure.

Roughly paralleling the results reported by Coke et al. (1978), who had found that 19 of 44 subjects (.43) volunteered to help Katie, we found that 61 of our 120 subjects (.51) volunteered some time. Given that almost half of our subjects did not help, scores on the continuous measure were badly skewed. Therefore, we adopted the dichotomous measure as our major index of helping, but as a check we also analyzed the scaled measure. The proportion of subjects volunteering to help Katie in each of the six cells is presented in Table 4; the means on the scaled measure are also shown.

Table 4

<table>
<thead>
<tr>
<th>Proportion of Subjects in Each Justification Condition Who Helped in Low- and High-Empathy Conditions: Study 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Justification condition</td>
</tr>
<tr>
<td>-------------------------</td>
</tr>
<tr>
<td>Replication condition</td>
</tr>
<tr>
<td>(no justification info)</td>
</tr>
<tr>
<td>Low justification for not helping</td>
</tr>
<tr>
<td>High justification for not helping</td>
</tr>
</tbody>
</table>

Note. N = 20 (10 men, 10 women) per cell. Means are for the scaled measure of helping (0 = no helping, 1 = 1 to 2 hr; 2 = 3 to 5 hr; 3 = 6 to 8 hr, and 4 = 9 to 10 hr).

Effects of the justification manipulation on the proportion of subjects helping in each cell of the 2 (low vs. high empathy) x 2 (low vs. high justification for not helping) design appear in the last two rows of Table 4. As can be seen, the pattern of helping across the four cells of the 2 x 2 design was consistent with predictions from the empathy-altruism hypothesis: lower helping in the low-empathy/high-justification cell than in the other three. A planned comparison testing this 1 versus 3 pattern was highly significant, \( \chi^2(1, N = 80) = 15.39, p < .001 \), and accounted for all reliable between-cell variance, residual \( \chi^2(2, N = 80) = 1.35 \). Pairwise comparisons revealed that the rate of helping in the low-empathy/high-justification cell was significantly lower than in each of the other three cells, all zs > 2.75, ps < .005, and there were no reliable differences among the other three (all zs < 1.0).

Analysis of the scaled measure of helping produced exactly the same pattern of significant effects, although the pattern was a little weaker than for the dichotomous measure, presumably

---

6 There was a significant effect of sex on reported concentration on Katie's feelings; women reported more concentration on her feelings (M = 7.60) than did men (M = 6.98), F(1, 108) = 4.31, p < .04. However, there were no interactions of sex with the experimental manipulations, all Fs(1, 108) < 1.70. The main effect for sex suggested that women either were more inclined than men to imagine the feelings of our female victim or were more inclined to present themselves as having imagined her feelings. In either case, given the absence of interactions, this main effect did not call into question the effectiveness of the listening-perspective instructions in focusing both men and women subjects' attention either away from or toward Katie's feelings about her plight.
due to the skew. For the scaled measure, the planned comparison testing the 1 versus 3 pattern predicted by the empathy-altruism hypothesis was highly significant, $F(1, 76) = 11.04$, $p < .001$, and accounted for all reliable between-cell variance, residual $F(2, 76) < 1.0$. Mean helping in the low-empathy/high-justification cell was significantly lower than in each of the other three cells, all $t_s(76) > 2.40$, $ps < .02$, and there were no reliable differences among the other three, all $t_s(76) < 1.0$.

These analyses indicated that subjects' helping responses conformed closely to the pattern predicted by the empathy-altruism hypothesis. There was no evidence of the significant effect of justification manipulation in the high-empathy condition that had been predicted by the empathy-specific punishment hypothesis.

Finally, some empirical evidence for the validity of our manipulation-check question concerning perceived obligation was provided by the correlation between responses to this question and helping. Across the entire design, subjects who reported feeling more strongly that students ought to help Katie were indeed somewhat more likely to help, $r_{pbis}(118) = .23$, $p < .02$, and to help more, $r(118) = .26$, $p < .01$. There were no reliable between-cell differences in these correlations, although they tended to be highest in the high-empathy/low-justification cell (both $r_s = .32$) and lowest in the low-empathy/high-justification cell ($r_s = .04$ and .14 for the dichotomous and scaled helping measures, respectively).

**Study 3: Justification for Not Helping Through Attributional Ambiguity**

Study 3 provided a generalized replication of the same Empathy × Justification design used in Study 2. In Study 3, however, the need situation, helping response, and both independent variables—empathy and justification for not helping—were operationalized differently. The person in need was a same-sex peer who had been (ostensibly) randomly assigned to receive a moderately uncomfortable electric shock for each error he or she made on a task. Subjects were to perform a different task; they would have a choice of two options and would not be punished for errors. Instead, for each correct response on Option A, they would receive one ticket for a raffle with a prize worth $30; for each correct response on Option B, they would reduce by one the shocks the peer was to receive. Spending time working on Option B rather than on Option A constituted helping. An attempt to manipulate empathy by varying similarity to the person in need proved ineffective; our similarity information had no reliable effect on either subjects' reports of empathy or their helping. Therefore, we relied on subjects' self-reports of empathy to define low- and high-empathy groups. As pointed out earlier, this measurement technique has proved as effective as experimental manipulation in operationalizing empathy as an independent variable and in producing the empathy-helping relationship.

Justification for not helping was manipulated by varying information about helping-irrelevant attributes of the two task options. In the low-justification condition both options involved either numbers or letters, and no information was given about other people's preferences. In the high-justification condition, the two options differed in that one involved numbers and the other letters. Subjects in this condition were told that most people prefer to work with the numbers (letters), whichever appeared on the nonhelpful Option A. (Pairing of the numbers and letters with the two options was counterbalanced.)

We reasoned that individuals who were given two different task options and were told that most people prefer one over the other would be provided with attributional ambiguity for their choice of that option (Snyder et al., 1979). They could attribute choosing to work on the nonhelpful option to the type of task (numbers or letters) rather than to selfishness, reducing anticipated self-punishment. Individuals for whom both options involved either numbers or letters could not justify choosing to work on Option A in this way.

**Predictions**

Predictions for Study 3 were essentially the same as for Study 2. The empathy-specific punishment hypothesis predicted less helping in the high-justification condition than in the low for both low- and high-empathy subjects. The empathy-altruism hypothesis predicted less helping in the high-justification condition than in the low for low-empathy subjects but little or no effect of the justification manipulation among high-empathy subjects. Across the four cells of the 2 (low vs. high empathy) × 2 (low vs. high justification) design, then, the pattern of results most consistent with the empathy-altruism hypothesis would be the same 1 versus 3 pattern as in Study 2; the rate of helping in the low-empathy/high-justification cell should be lower than the rate in the other three cells.

**Method**

**Subjects**

For Study 3, 88 students (48 men, 40 women) in an introductory psychology course at the University of Kansas served as subjects, partially fulfilling a course requirement. Within sex, a randomized-block procedure was used to assign subjects to the two justification conditions, 43 (23 men, 20 women) to the low-justification condition and 45 (25 men, 20 women) to the high. (A blocking error caused the unequal number of men in the cells.)

**Procedure**

Because the procedure was very similar to the one used in Study 1, we discuss in detail only the new features. As before, subjects were told that the study concerned "how a variety of task characteristics and outcome consequences affect people's task performance and attitudes." This time, however, the task characteristics being examined were (a) whether one has a choice of tasks on which to work; (b) for whom one's performance has consequences—for self only, for another person only, or both; (c) the kind of outcome consequences—positive, negative, or neutral; and (d) complexity of the task situation—simple (varying only in whether the consequences for self are positive, negative, or neutral) or complex (varying on choice, type of consequences, and for whom).

---

7 Thirteen additional students (5 men, 8 women) were excluded from the sample and replaced because of suspicion. The suspicion rate did not differ reliably between the justification conditions, $\chi^2(1, N = 101) = 0.11$, and analyses the same as those reported here but including suspicious individuals produced the same pattern of significant effects.
Positive consequences were tickets in a raffle for a $30 gift certificate; negative consequences were mild but uncomfortable electric shocks; and neutral consequences were just information.

**Condition 9:** choice of tasks, one benefiting self and one benefiting another person. All subjects were informed that they had (ostensibly randomly) been assigned to Condition 9, a complex task situation in which they would have a choice between two task options, both with positive consequences. They could divide their time between the two options in any way they chose: Option A would have positive consequences for them, and Option B would have positive consequences for the person assigned to the simple task situation. A sheet provided more detailed information about Condition 9:

Since you have been assigned to the positive consequences condition, you will be given a raffle ticket for every point you receive on Option A. Depending on the condition initially assigned to the person in the simple task situation—negative or positive—he or she will be given fewer shocks or more raffle tickets for each point you receive on Option B.

The sheet also informed subjects that before performing their task they would get to know something about the person assigned to the simple task situation (Brian for men, Janet for women) by listening as he (she) talked to them briefly over an audio intercom.

Once subjects read this sheet, the experimenter told them that Brian (Janet) had been initially assigned to the negative consequences condition; he (she) would receive an electric shock for each error on his (her) task. Each point the subject received on Option B would eliminate one of those shocks. Brian (Janet) was not aware of the subject's opportunity to help him (her) avoid the shocks.

**Empathic reaction to learning of the other person's need.** After being given this information, subjects were left alone to listen to Brian (Janet) over the intercom. As before, what they heard was actually a prerecorded tape on which Brian (Janet) confessed to having concern about their shock. Depending on the condition, the experimenter repeated the following sentence: "You can work on whichever option you wish." If Option A involved letters, the words numbers and letters were reversed. The experimenter was blind to the subject's justification condition until just prior to presenting the options.

**Dependent measure: choosing to reduce Brian's (Janet's) shocks by working on Option B.** The experimenter then started the timer and left the subjects alone for 120 s to perform the task. The dependent measure of helping was the ratio of the number of correct combinations subjects circled on Option B divided by the total number they circled.

**Debriefing.** After subjects completed the task, they were carefully debriefed. They were also informed that a raffle for a $30 gift certificate would actually be held and that they would receive 10 tickets regardless of their task performance. Following debriefing, subjects were thanked for their participation and excused. Once all 88 subjects had been run, the raffle was held and the gift certificate awarded to the winner.

**Results and Discussion**

**Empathic Response to Brian's (Janet's) Need**

To provide an index of empathic reaction to Brian's (Janet's) need, responses to the five adjectives on the emotional response questionnaire that had been found in previous research to reflect empathy were averaged to form an index of empathic emotion (Cronbach's alpha = .87). Scores on this empathy index did not differ across justification conditions, t(84) < 1.0 (overall M = 3.88). We then performed a median split to identify low- and high-empathy subjects (MdN = 4.01). Crossing this low- versus high-empathy classification with the justification manipulation produced the Empathy × Justification design needed to test the empathy-specific punishment and empathy-altruism hypotheses.

**Effect of Justification for Not Helping on the Helping of Low- and High-Empathy Subjects**

The measure of helping was the ratio of number of combinations correctly circled on Option B—the option that would reduce the number of shocks Brian (Janet) would receive—divided by total number of combinations correctly circled on both options. This ratio could range in value from 0 if all of the combinations circled were on Option A, the option that would earn tickets in the $30 raffle, to 1.0 if all of the combinations circled were on Option B. The mean helping response on this ratio scale for subjects in each cell of Study 3 is presented in Table 5.

As can be seen, the pattern of helping across the four cells of the 2 × 2 design was highly consistent with the predictions of the empathy-altruism hypothesis: lower helping in the low-empathy/high-justification cell than in the other three. A planned comparison testing this 1 versus 3 pattern was highly significant.

---

**Footnote:** As in Study 1, we found a significant sex difference on the empathy index; women reported more empathy for Janet (M = 4.56) than men reported for Brian (M = 3.47), F(1, 84) = 9.17, p < .003. This sex difference could not, however, account for the observed pattern of helping; the same pattern of significant results was found when the median split on empathy was performed separately within sex.

We did not attempt to create an index of personal distress in this study because, as in Study 1, our distress adjectives seemed to measure an other-oriented, empathic response in this research paradigm.
The justification cell was significantly lower than in each of the other conditions, as predicted by the empathy-specific punishment hypothesis. However, the performance standard on the digit scan recognition task was significantly higher than the performance standard to qualify to help is easy enough that most people do not meet it, they should not feel that they can gracefully fail to qualify. Therefore, to avoid self-recrimination, they should try harder at the qualifying task. They should also try harder at the qualifying task when the standard is difficult than when it is easy. This is because only by offering to help and qualifying can they reach their altruistic goal of reducing the other person's suffering.

To test the relative merits of the empathy-specific punishment and the empathy-altruism hypotheses using this logic, we needed a research paradigm in which (a) helping is personally costly, (b) the empathy-helping relationship is known to occur, and (c) it is plausible to introduce a qualifying task for helping. The shock paradigm developed by Batson et al. (1981; see also Batson et al., 1983, 1986) seemed to meet these requirements.

**Introducing a Qualifying Task Into the Shock Paradigm**

In this paradigm, female subjects watched over closed-circuit TV (actually a videotape) as a young woman, Elaine, appeared to receive electric shocks while performing a digit recall task. Elaine's reactions made it clear that she was finding the shocks highly uncomfortable. After Elaine completed only two of her scheduled 10 digit recall trials, subjects were unexpectedly given a chance to relieve her distress by taking the remaining eight trials, and the shocks, in her stead.

We introduced a qualifying task into this shock paradigm by explaining that we were interested in studying auditory numeric facility, measured by the digit recall task, of only those individuals who displayed a certain level of visual numeric facility, which was measured by a digit scan recognition task. Thus, even if subjects were willing to help Elaine by taking her place, they would be allowed to do so only if they met a certain performance standard on the digit scan recognition task.

**Operationalizing Empathy Through Self-Reported Predominant Emotional Response**

As in Study 1, we operationalized empathy as an independent variable through measurement rather than manipulation. Batson et al. (1983) pointed out that when assessing self-reported empathy in a high-impact, proximate situation such as watching a peer receive shocks, it is important to differentiate empathic feelings from feelings of personal distress. Both personal distress and empathy are possible emotional reactions to witnessing another person suffer, and they often occur together. Feelings of personal distress are more self-oriented; they are the upshot and discomfort one feels as a result of witnessing another person's suffering. Feelings of empathy, on the other hand, are more other-oriented; they are the sympathy and compassion one feels for the suffering person. There are both theoretical and empirical grounds for believing that these two emotional reactions to witnessing another person suffer have different motivational consequences: Whereas the motivation to help associated with feeling empathy may be altruistic, the motivation to help associated with personal distress seems clearly egoistic (see Batson, 1987, for a discussion of the distinction between personal distress and empathy and a review of the evidence).

In order to assess the nature of the motivation to help associated with self-reported empathy in a situation in which both distress and empathy are felt, Batson et al. (1983) recommended the use of an index of predominant emotional response, created by subtracting each subject's score on an index of personal distress from his or her score on an index of empathy. This index of predominant emotional response should reveal...
flect the relative strength of competing motives associated with distress and empathy. If, for example, an individual experiencing a high degree of empathy were experiencing an even higher degree of personal distress, motivation associated with the latter emotion should dominate.

Following the recommendation of Batson et al. (1983), we operationalized empathy as an independent variable by performing a median split on subjects' scores on an index of predominant emotional response. This split produced two groups, one experiencing a relative predominance of personal distress and the other a relative predominance of empathy.

**Ease of Escape**

One final aspect of the shock paradigm needed to be considered. In previous studies using this paradigm, ease of physical escape from Elaine's suffering had been manipulated. This was done by leading some subjects to believe that even if they chose not to help, they would not watch Elaine take the remaining shocks (easy escape); other subjects believed that if they chose not to help, they would watch Elaine's remaining eight trials (difficult escape). Because we wished to examine the effects of difficulty of the performance standard both on whether subjects would offer to help and, if they offered, on their performance on the qualifying task, we decided that it was best to lead all subjects to believe that if they chose not to help, they would watch Elaine's remaining eight trials. In previous studies, this difficult-escape information had produced a high rate of helping among subjects reporting a predominance of distress, as well as among those reporting a predominance of empathy.

**Predictions**

The empathy-specific punishment hypothesis predicted that the difficulty of the qualifying standard would have much the same effect on the behavior of subjects feeling a relative predominance of distress and those feeling a relative predominance of empathy because the motivation to help of subjects in both groups would be egoistic: Either the rate of helping or performance on the qualifying task, or both, should be lower when the qualifying standard was difficult than when it was easy. The empathy-altruism hypothesis made the same prediction for difficult qualifying-standard condition. More subjects were assigned to the difficult-standard condition because we anticipated a lower rate of helping in that condition (at least among subjects reporting a relative predominance of distress), and we wished to ensure that even in this condition we had enough subjects offering help that we could meaningfully analyze performance on the qualifying task. Nine additional students (5 in the easy- and 4 in the difficult-standard condition) were excluded from the design and replaced because they expressed suspicion that the confederate was not actually receiving shocks.

**Procedure**

Subjects were run individually. Because the experimental procedure was much the same as that used by Batson et al. (1986), we describe in detail only the new features.

**Initial administration of the digit scan recognition task.** When subjects arrived, they were first asked to complete a digit scan recognition task, ostensibly while waiting for the other participant, Elaine (actually fictitious). Subjects were told that the purpose of the digit scan task would become clear later. The task was the same used in Study 1; it involved working for 120 s from left to right, by line down a page filled with randomly ordered digits, circling all occurrences of the combinations 13 and 47. Unknown to subjects, this was exactly the same task they would later be asked to perform in order to qualify to help Elaine. This initial administration served to familiarize them with the task and provided a context for introducing the qualifying-standard manipulation.

**Manipulation of qualifying-standard difficulty:** After performing the task, subjects read an introduction that described the study as being concerned with the effects of aversive conditions on task performance and impression projection. Subjects learned that one participant would serve as a worker, performing a numeric recall task while receiving mild electric shocks at random. The other participant would serve as the observer, watching over closed-circuit TV while the worker performed the recall task. The observer's job was to form and report a general impression of the worker.

The introduction also explained that the numeric recall task performed by the worker involved only one aspect of numeric facility: auditory recall of number sequences. Because it was necessary that anyone serving as the worker have numeric facility, everyone first completed the digit scan task, which was a visual measure of numeric facility. The introduction explained that, "Only individuals who display numeric facility by performing at a certain standard on the digit scan task will be able to be the worker on the numeric recall task."

The experimental manipulation of qualifying-standard difficulty was introduced in the last paragraph of the introduction. In the easy qualifying-standard condition, subjects read the following: "A moderately stringent standard has been adopted for performance on the digit scan task. On the average, about 7 of 10 college students meet the standard." In the difficult qualifying-standard condition, subjects read the following: "An extremely difficult standard has been adopted for performance on the digit scan task. On the average, only 1 of 5 college students meet the standard, so do not be surprised or disturbed if you do not." In both conditions subjects were informed that "numeric facility is not generally associated with intelligence or mathematical ability. It is simply a capacity that some people display and some do not." The experimenter did not know which version of the introduction a subject received and so remained blind to the subject's experimental condition.

**Observer role.** Once subjects had read the introduction and agreed to participate in the study, they were all (ostensibly randomly) assigned to the observer role; Elaine was assigned to the worker role. Noting that Elaine had met the standard on the digit scan task, the experimenter commented, "Since you'll be the observer, we don't need your digit scan task; I didn't even score it," and tossed it in the wastebasket.
Self-reported emotional response to Elaine’s distress. Subjects were then left alone to observe Elaine over closed-circuit TV as she performed 10 digit recall trials while receiving the shocks. By the end of the second trial, the shocks seemed to be hurting Elaine so much that Marsha, the assistant administering the shocks, interrupted the procedure and asked whether Elaine was all right. Elaine hesitantly said yes, then asked for a glass of water before going on. Marsha agreed and left to get the water.

During this break, the experimenter spoke to subjects over an audio intercom, explaining that there was going to be a short delay, they should complete the first impression questionnaire. This questionnaire, called the emotional response questionnaire, consisted of a list of 28 adjectives describing emotions. Subjects were asked to indicate on 7-point scales (1 = not at all; 7 = extremely) how much they were experiencing each emotion as a result of observing the worker. The list of emotions included adjectives that have been found in past research (Batson, 1987; Batson et al., 1983; Coke et al., 1978; Tol & Batson, 1982) to reflect the two distinct vicarious emotions: personal distress and empathy.

After about 90 s, Marsha returned and asked Elaine whether she had ever before been bothered by electric shock. Elaine then confessed to having had a traumatic experience with shock as a child. (This information was provided in order to ensure that subjects would consider Elaine’s extreme reaction to the shocks atypical and would not expect to find the shocks as unpleasant if they chose to take her place.) Hearing this, Marsha suggested that the experiment be stopped, but Elaine expressed a firm desire to continue: “I started; I want to finish. I’ll go on... I know your experiment is important, and I want to do it.” At this point, Marsha had an idea: The observer was also an introductory psychology student; maybe she would be willing to help Elaine out by taking her place as the observer. Elaine, with a mixture of reluctance and relief, agreed to allow the observer to be asked. The tape ended with Marsha turning off the video equipment to go check with the experimenter about this possibility.

Helping opportunity. About 20 s later, the experimenter’s voice, sounding somewhat agitated, came over the intercom and outlined the subject’s options: either to remain as the observer, watching Elaine’s behavior, or to help Elaine by taking her place “doing the recall task and receiving the shocks.” The experimenter then added the following:

Oh, I almost forgot. Even if you volunteer to take Elaine’s place, you’ll be able to do so only if you meet the qualifying standard on the digit scan task—you know, the number-circling task on which, on the average, [easy qualifying-standard condition: 7 of 10 college students qualify] [difficult qualifying-standard condition: only 1 of 5 college students qualify]. Oh, I guess I messed up your earlier try at that, so I’ll need you to do that task again to see if you qualify.

Finally, the experimenter emphasized, “whichever you want to do is fine.” This communication was prerecorded, one tape for each qualifying-standard condition. At this point in the procedure, the experimenter simply selected the appropriate tape and played it over the intercom.

Helping response. The subject was left alone for about 30 s after the tape ended, giving her time to decide. The experimenter then entered the observation room and asked for the subject’s decision, which was the measure of helping.

Performing the qualifying task. If the subject decided to help Elaine, the experimenter placed a new copy of the digit scan task, identical to the one completed earlier, in front of the subject and left her alone for 120 s to work on it. After completing this task, subjects were asked to fill out a second impression questionnaire and then a digit scan task questionnaire while the experimenter went to score their digit scan task.

If the subject decided not to help, she did not perform the qualifying task or complete the digit scan task questionnaire. She simply completed the second impression questionnaire while the experimenter went to tell Marsha what had been decided.

Ancillary measures. The second impression questionnaire assessed subjects’ perceptions of the worker, including how uncomfortable the shocks were for her. This assessment provided a measure of the perceived severity of Elaine’s need. The first item on the digit scan task questionnaire provided a check on the effectiveness of the performance standard manipulation; it asked subjects how difficult it would be to meet the qualifying standard.

Debriefing. After subjects completed these questionnaires, the experimenter returned and informed them that because of lack of time it had been decided that the session would not continue; neither Elaine nor they would do any more recall trials or receive any more shocks. The experimenter then expressed a desire to talk with subjects for a few minutes, getting their reactions to the study. All subjects readily agreed. After discussing their reactions, subjects were carefully debriefed, thanked for their participation, and excused.

Results and Discussion

Effectiveness of the Qualifying-Standard Manipulation

The manipulation–check item on the digit scan task questionnaire suggested that the qualifying-standard manipulation was successful. An ANOVA indicated that the only reliable effect on responses to the question “How difficult do you think it will be for you to meet the performance standard on the digit scan task?” (1 = not at all difficult, 9 = extremely difficult) was a main effect for the qualifying-standard manipulation, F(1, 32) = 5.36, p < .03. Subjects in the difficult qualifying-standard condition rated the standard as being more difficult (M = 6.25) than did subjects in the easy qualifying-standard condition (M = 4.90).

Perceptions of Elaine’s Need

Subjects in both experimental conditions perceived Elaine to be in considerable need. The overall mean response to the question of how uncomfortable the shocks were for the worker, rated on a 7-point scale (1 = not at all, 7 = extremely), was 6.50. Mean response did not differ reliably across conditions, t(58) < 1.0.

Index of Predominant Emotional Response: Personal Distress Versus Empathy

Each subject’s responses to eight adjectives on the emotional response questionnaire found in past research to reflect feelings of personal distress (alarmed, grieved, troubled, distressed, upset, disturbed, worried, and perturbed) were averaged in order to form an index of self-reported personal distress at watching Elaine suffer (Cronbach’s alpha = .94); responses to the six adjectives previously found to reflect feelings of empathy (sympathetic, moved, compassionate, warm, soft-hearted, and tender) were averaged to form an index of self-reported empathy (Cronbach’s alpha = .89). As in previous studies, observing Elaine evoked relatively high levels of both distress and empathy; the mean on the 7-point response scale was 5.34 for the distress index and 4.58 for the empathy index. For neither index was
there a reliable effect for the qualifying-standard manipulation, both \( t(58) < 1.10 \).

Following the same procedure as Batson et al. (1983), we created an index of predominant emotional response by subtracting each subject’s score on the distress index from her score on the empathy index. We then performed a median split on this index of predominant emotional response (\( M_{dn} = -.20 \)), dividing our subjects into those who felt a relative predominance of distress and those who felt a relative predominance of empathy. This median split, combined with the experimental manipulation, produced the 2 (easy vs. difficult qualifying standard) \( \times \) 2 (predominant distress vs. empathy) design needed to test the predictions of the empathy-specific punishment and the empathy-altruism hypotheses.

**Rate of Helping in Each Condition of the 2 \( \times \) 2 Design**

Of the 60 subjects, 36 offered to help Elaine by taking the shocks in her stead. The proportion of subjects who offered help in each cell of the 2 \( \times \) 2 design is presented in Table 6. As in Studies 2 and 3, the pattern of helping across the four cells was highly consistent with the predictions of the empathy-altruism hypothesis: less helping in the difficult standard/distress cell (.28) than in the other three cells (.65 or higher). A planned comparison testing this 1 versus 3 pattern was highly significant, \( \chi^2(1, N = 60) = 10.47 \), \( p < .001 \), and accounted for all reliable between-cell variance; residual \( \chi^2(2, N = 60) = 1.84 \) (analyses using normal approximation based on arc sine transformation; see Study 2). Pairwise comparisons revealed that the proportion of subjects offering help in the difficult standard/distress cell was significantly lower than in each of the other three cells, all \( z > 2.04 \), \( p < .025 \), and there were no reliable differences among the other three (all \( z < 1.35 \)).

**Performance on the Qualifying Task**

We then examined performance on the qualifying task by those subjects who agreed to help. Across the four cells of the 2 \( \times \) 2 design, the empathy-specific punishment hypothesis predicted a main effect for difficulty of the qualifying standard; performance should be worse when the standard was difficult. In contrast, the empathy-altruism hypothesis predicted that performance by prospective helpers reporting a predominance of empathy would be the same or better when the standard was difficult. Rather than predicting a main effect for difficulty of the qualifying standard, the empathy-altruism hypothesis predicted a Difficulty of Standard \( \times \) Predominant Emotional Response interaction.

**Measuring performance**. The number of combinations correctly circled on the second digit scan task served as our performance measure. Because only the 36 subjects who offered to help performed the qualifying task, we created a new median split on the index of predominant emotion for these 36 helpers (\( M_{dn} = .00 \)), providing a balanced 2 (easy vs. difficult standard) \( \times \) 2 (predominant distress vs. empathy) design. (To ensure that this new split did not introduce some artifact, we also examined performance using the old split. The pattern of means was the same reported here, as was the pattern of significant effects.)

**Performance on the digit scan task**. The mean number of combinations circled correctly by subjects in each cell of the 2 \( \times \) 2 design are presented in Table 7. An ANOVA revealed a marginally significant main effect for predominant emotional response, \( F(1, 32) = 3.02 \), \( p < .10 \); subjects reporting a predominance of empathy circled more combinations (\( M = 11.28 \)) than did subjects reporting a predominance of personal distress (\( M = 8.25 \)). This main effect was, however, qualified by a highly significant Difficulty of Standard \( \times \) Predominant Emotional Response interaction, \( F(1, 32) = 10.17 \), \( p < .003 \). There was no evidence of the main effect for the qualifying-standard manipulation predicted by the empathy-specific punishment hypothesis, \( F(1, 32) < 1.0 \).

Pairwise comparisons indicated that the interaction was the same as that predicted by the empathy-altruism hypothesis. Subjects reporting a predominance of distress circled significantly fewer combinations when the qualifying standard was difficult (\( M = 8.25 \)) than when it was easy (\( M = 11.30 \)), \( t(32) = 2.24 \), \( p < .02 \), one-tailed. Subjects reporting a predominance

\[ \text{Table 7} \]

**Performance on Qualifying Task by Subjects Offering Help in Each Cell of 2 \( \times \) 2 Design: Study 4**

<table>
<thead>
<tr>
<th>Qualifying standard</th>
<th>Predominant emotional response</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Distress</td>
</tr>
<tr>
<td>Easy</td>
<td>.73</td>
</tr>
<tr>
<td>( M )</td>
<td>10</td>
</tr>
<tr>
<td>( n )</td>
<td>14</td>
</tr>
<tr>
<td>Difficult</td>
<td>.28</td>
</tr>
<tr>
<td>( M )</td>
<td>8</td>
</tr>
<tr>
<td>( n )</td>
<td>17</td>
</tr>
</tbody>
</table>

Note. Performance was measured by the number of combinations correctly circled. Cell means not sharing a common subscript differ, \( p < .05 \), by a \( t \) test.

9 A varimax-rotated principal-components analysis of all 14 of these emotional response adjectives produced a two-component solution similar to the solutions found in previous studies (see Batson, 1987); all 8 distress adjectives loaded above .65 on one component, and all 6 empathy adjectives loaded above .65 on the other orthogonal component.
of empathy circled significantly more combinations when the qualifying standard was difficult ($M = 13.00$) than when it was easy ($M = 9.90$), $t(32) = 2.27$, $p < .02$, one-tailed. The performance difference in the difficult qualifying-standard condition between subjects reporting a predominance of distress and those reporting a predominance of empathy was highly reliable, $t(32) = 3.30$, $p < .001$, one-tailed.

Implications of Studies 2 Through 4

Providing justification for not helping—whether by the inaction of others (Study 2), by introducing attributional ambiguity (Study 3), or by making it difficult to qualify to help (Study 4)—had a dramatic effect on the helping of low-empathy subjects; it had little effect on the helping of high-empathy subjects. This pattern of results, which was highly consistent across studies, suggested that although the helping of low-empathy subjects was motivated at least partly by a desire to avoid the self-punishment associated with a failure to do the right thing, the helping of high-empathy subjects was not. The relatively high rate of helping by high-empathy subjects, even when justification for not helping was high, was precisely what we would expect if, as claimed by the empathy—altruism hypothesis, feeling empathy for the person in need evoked altruistic motivation to have that person's need reduced. It was not what we would expect if, as claimed by the empathy-specific punishment hypothesis, feeling empathy evoked increased egoistic motivation to avoid anticipated self-punishment.

Did Our Justification Manipulations Really Reduce Empathy-Specific Punishment?

Even though the results of these three studies certainly seem to count against the empathy-specific punishment hypothesis, they do not entirely rule it out. Our manipulations of justification for not helping were indirect, so we cannot be absolutely certain that they reduced anticipated empathy-specific self-punishment. The finding that our justification manipulations had a significant effect on helping of low-empathy subjects, coupled with the finding by Fultz et al. (1986) of no reduction even in a low-empathy condition for a powerful manipulation of anticipated social punishment, suggests that our manipulations did reduce anticipated self-punishment for a failure to help. But to press the point, perhaps the information that we provided to justify not helping was only effective in reducing general and not empathy-specific self-punishments. Perhaps anticipation of the latter is so well internalized that it is impervious to information about the inactivity of other potential helpers, the helping-irrelevant attributes of the helping task, or the difficulty of qualifying. If this were true, could the empathy-specific punishment hypothesis not still account for the observed pattern of results: substantially reduced helping in the high-justification condition for low- but not for high-empathy subjects?

Indeed it could, but the relevant evidence suggests that our justification manipulations were not ineffective in reducing anticipated empathy-specific punishment. Recall that on the perception of obligation measure in Study 2 there was evidence that learning most peers had not helped reduced the perception of obligation to help Katie as much in the high-empathy condition as in the low. If high-empathy subjects are impervious to justification information, there should have been no change in their perception of obligation. Recall also that the perception of obligation by subjects in the replication cells of Study 2, who were provided with no information about the helping of others, was essentially the same in the high-empathy condition as in the low, and in neither replication condition did perceived obligation differ reliably from the perception in the high-justification condition. On our obligation measure, then, there was no evidence of the increased sense of obligation associated with feeling empathy claimed by the empathy-specific punishment hypothesis.

But to persist, perhaps our obligation measure was flawed. Perhaps high-empathy subjects in the replication condition perceived a change in their own personal obligation to help Katie that was not picked up by our measure, which (to be less reactive) had asked whether “students” ought to help. Perhaps the effect of the justification manipulation on perceived obligation in the high-empathy condition reflected a change in perception of what other students ought to do, but not what the empathically aroused subject felt he or she ought to do. This may be true, but recall that responses on the perception of obligation measure were positively correlated with subjects’ own helping, suggesting that it reflected implications for their own behavior.

Could the Lack of Difference Among High-Empathy Subjects Be Due to a Ceiling Effect?

Another possibility is that the lack of difference in helping between the high- and low-justification conditions for high-empathy subjects was caused by a ceiling effect. Recall that the obligation measure indicated that the information about other helpers in Study 2 might have increased perceived obligation to help in the low-justification condition at least as much as it decreased the perceived obligation in the high-justification condition. This might have been because, as suggested in the replication condition, the perceived obligation to help was not especially high even in the absence of justifying information. Whatever the reason, the increase in obligation raises the possibility that in Study 2 a ceiling effect might have prevented us from observing an increase in helping among high-empathy/low-justification subjects. The proportion of helping in the high-empathy/replication cell was .70. Even though this proportion is not close to the absolute ceiling of 1.00, perhaps it is close to a functional ceiling. Perhaps helping in the high-empathy/low-justification cell would have been significantly higher than in the high-empathy/high-justification cell had this functional ceiling not prevented it from rising. A similar argument can be made concerning helping in Study 4.

A close look at our data renders this ceiling-effect explanation implausible. In Study 2 we found exactly the same pattern of significant effects on the scaled measure of helping (on which the mean in the high-empathy/replication cell was 1.20 on a 0 to 4 scale) as we found for the dichotomous measure based on proportions. In Study 3, we again found the same pattern of significant effects on a scaled measure, and in that study, unlike Study 2, there was no information provided in the low-justification condition to increase the obligation to help. Given the same pattern of helping across measures and across studies, it
is not likely that the absence of an effect of our justification manipulations on the helping of high-empathy subjects was caused by a ceiling effect.

Instead, we believe that the results of these three studies seriously challenge the anticipated self-punishment version of the empathy-specific punishment hypothesis. Still, in a final attempt to find evidence for this hypothesis, and for the empathy-specific reward hypothesis, we examined the goal-relevant cognitions associated with empathy-induced helping.

GOAL-RELEVANT COGNITIONS ASSOCIATED WITH EMPATHY-INDUCED HELPING

The empathy-specific reward, empathy-specific punishment, and empathy-altruism hypotheses each postulate a different motive underlying the helping associated with feeling empathy. These different motives have different goals: seeking rewards, avoiding punishments, and relieving the victim's need, respectively. Because each hypothesis assumes that the empathically aroused individual, when considering whether to help, has in mind one of these goals, each assumes that cognitions relevant to that goal should be especially salient. Determining whether reward-relevant, punishment-relevant, or victim-relevant cognitions are associated with empathy-induced helping should provide some evidence as to the goal of empathically aroused helpers. A Stroop task seemed to be one means of determining the salient goal-relevant cognitions.

Assessing Goal-Relevant Cognitions Using a Stroop Task

A Stroop task (Stroop, 1938) involves having subjects name as quickly as possible the color of the ink in which a word or other visual stimulus appears. As Geller and Shaver (1976, p. 101) observed: "In general, it appears that latency of color naming for a particular word will increase whenever a subject has been thinking about something related to that word" (see also Schadler & Thissen, 1981; Warren, 1974). This Stroop procedure seemed capable of providing a measure of the nature of the motivation underlying the empathy--helping relationship, one that was relatively nonreactive and did not rely on subjects' conscious awareness of or willingness to report their goals.

Predictions

If the increased helping associated with empathic emotion is motivated by the desire to reach a given goal, then this helping should be positively correlated with the latency to name the color of words relevant to that goal. This observation leads to three distinct predictions: (a) If the motivation to help associated with feeling empathy is directed toward the egoistic goal of obtaining social and self-rewards, as the empathy-specific reward hypothesis claims, then the increased helping associated with empathic emotion should be positively correlated with color-naming latency for reward-relevant words. (b) If the motivation to help associated with feeling empathy is directed toward the egoistic goal of avoiding social and self-punishments, as the empathy-specific punishment hypothesis claims, then the increased helping associated with empathic emotion should be positively correlated with color-naming latency for punishment-relevant words. (c) If the motivation to help associated with feeling empathy is directed toward the altruistic goal of reducing the victim's suffering, as the empathy--altruism hypothesis claims, then the increased helping associated with empathic emotion should be positively correlated with color-naming latency for victim-relevant words. Although distinct, these three predictions are not mutually exclusive; correlations may be found supporting any or all.

If we could assume a one-to-one correspondence between either manipulated or measured empathic emotion and increased reward-seeking, punishment-avoiding, or altruistic motivation, then we could predict a direct association between increased empathy and increased color-naming latency for words of one or more of the three types. It seemed likely, however, that situational factors would make it impossible for some empathically aroused individuals to help. This could inhibit their motivation and the salience of related goal-relevant cognitions (see Brehm, Wright, Solomon, Silka, & Greenberg, 1983). Therefore, it seemed best to direct our attention to the more precise relationship predicted by the three hypotheses: the relationship between increased helping associated with feeling empathy and increased color-naming latency for words of one or more of the three types. If empathically induced helping is directed toward gaining rewards, avoiding punishments, or reducing the victim's suffering, then the color-naming latency for words relevant to one or more of these goals should mediate the relationship between feeling empathy and helping (Baron & Kenny, 1986).

A Preliminary Attempt to Test the Predictions

C. D. Batson, J. Orendain, D. Shetrompf, and M. L. Templeton made a preliminary attempt to test these predictions. They conducted a pilot study in which 23 female undergraduates read about Sandy, a lonely, disadvantaged 12-year-old girl seeking a surrogate grandparent to be an adult friend and guide. After reading about Sandy, subjects were informed that they would soon be given a chance to volunteer time to write letters to prospective "grandparents" on her behalf. First, however, ostensibly as a baseline control for a measure of cognitive reactions to the information about Sandy, subjects were asked to name as quickly as possible the color (red, blue, green, or brown) in which each word in a series appeared. Some of the words were reward relevant (good, merit, honor, and praise); some were punishment relevant (duty, guilt, shame, and should); some were victim relevant (hope, child, needy, and friend); and some were neutral (left, rapid, large, and breath). Using a millisecond timer and voice-operated relay, it was possible to assess color-naming latencies for the different types of words by repeated measures within subjects (see Schadler & Thissen, 1981; Warren, 1974) rather than by the more gross between-group comparisons used by Stroop (1938) and Geller and Shaver (1976). After completing this Stroop task, helping was measured by the amount of time, if any, subjects volunteered to spend writing letters for Sandy.

In an attempt to manipulate empathy, written instructions directed some subjects to focus on technical aspects of the information about Sandy (low empathy) and others to imagine how
she felt (high empathy), but this manipulation proved unsuccess-
ful. The amount of help offered did not differ reliably be-
tween the low- and high-empathy conditions, t(21) < 1.0, and
in debriefing, subjects assigned to the low-empathy condition
consistently reported that they had been unable to keep from
imagining how Sandy felt. Because it appeared that all subjects
either were placed or placed themselves in a high-empathy con-
dition, correlations between the amount of helping and color-
naming latencies were computed for all 23 subjects combined.

These correlations revealed a significant positive relationship
between amount of help volunteered and color-naming latency
for the victim-relevant words (averaged and adjusted for indi-
vidual differences in reaction time by subtracting the average
latency to the neutral words), r(21) = .53, p < .004, one-tailed.
Correlations between amount of helping and adjusted average
latencies for the reward-relevant and punishment-relevant
words were −.06 and −.10, respectively.

These correlations patterned as predicted by the empathy-
 altruism hypothesis, but not as predicted by either the empa-
thy-specific reward or the empathy-specific punishment hy-
thesis. However, because in this pilot study the empathy
manipulation failed and all subjects were assumed to be experi-
encing a high degree of empathy, results were not conclusive.

Study 5: Empathy, Helping, and a Stroop Task

Because the results were not conclusive, we conducted a new
study, using a need situation and empathy manipulation known
to produce the customary empathy–helping relationship. The
need situation and empathy manipulation we used were the
same as the ones used in Study 2 (see also Coke et al., 1978), in
which subjects adopted a particular perspective while listening
to a (bogus) radio broadcast that informed them of the need of
another undergraduate, Katie Banks.

Method

Subjects

Forty-eight female introductory psychology students at the University
of Kansas served as subjects, partially fulfilling a course require-
ment. By use of a randomized-block procedure, 24 subjects were assigned
to each of the two empathy conditions (low and high), and within these
conditions, two female experimenters each ran 3 subjects with each of
four Stroop slide sets (used to counterbalance words with colors and
order of presentation). Type of word (reward relevant, punishment rele-
vant, victim relevant, and neutral) was a repeated measures factor.10

Materials and Equipment

We used 16 stimulus words in the Stroop procedure, 4 of each type. Reward-relevant words were nice, proud, honor, and praise; punishment-relevant words were duty, guilt, shame, and obligé; victim-relevant words were loss, needy, adopt, and tragic; and neutral words were pair, clean, extra, and smooth. Across type, words were matched for length (one 4-letter word, two 5-letter words, and one 6-letter word) and approximate frequency of appearance in American English (Kucera & Francis, 1967).

To assess the appropriateness of the words selected to represent each
type, five independent judges first listened to the tape presenting Katie
Banks’s need, then classified the 16 words into the 4 types. Reward-
relevant words were correctly classified 80% of the time, punish-
ment-relevant words 100% of the time, victim-relevant words 100% of the
time, and neutral words 85% of the time. These results suggested that
the words were adequate representatives of the types.

Sixty-four Stroop slides were prepared, 1 of each of the 16 stimulus
words in each of four colors (red, blue, green, and brown) against a black
field. These slides were then arranged into four slide sets. Each word
appeared in each slide set, but in a different color. Moreover, within each
slide set, a word of each type and a word in each color appeared in the
first 4 words presented, also in the next 4, and so on, with no repetition
of specific word–color pairing or order across slide sets.

In each slide set, the 16 stimulus words were preceded by five buffer
slides. The first 4 presented a row of Xs in one of the four colors; the
fifth presented a neutral word, added because pretesting had revealed
an atypically long latency on the first word following the Xs. Opaque
slides were alternated with the slides of Xs and words to provide a dark
screen during the 5-s intertrial interval.

The slides were shown on a square 28-cm rear-projection screen
placed approximately 80 cm in front of subjects; letters of the projected
words were approximately 1 cm high and 0.5 cm wide. When each slide
was shown, a light-sensitive phototransistor attached to the projector
lens started a digital millisecond timer. The timer was stopped by the
electrical impulse from a Grayson-Stadler E7300A-1 voice-operated re-
delay, which was activated by input from a throat microphone as subjects
named the color in which the Xs or word was printed.

Procedure

The procedure was similar to that used by Coke et al. (1978), except
for the inclusion of the Stroop task and an emotional response question-
aire to measure self-reported empathy and the omission of the misat-
tribution manipulation. Therefore, we describe in detail only these new
aspects. The Stroop task was presented to subjects in a written introduc-
tion as a reaction time measure of the effect of the pilot radio broadcasts
on thoughts, and the emotional response questionnaire was presented as
a measure of the effect of the broadcasts on feelings.

The reaction time (Stroop) task. After the introduction, subjects read a
written description of the reaction time measure. This measure in-
volved looking at a series of words after each broadcast; some of the
words would be relevant to possible thoughts after hearing the broad-
cast, others would not. Different words would appear in different colors.
For each word, participants assigned to the “test group” would say as
quickly as possible whether it was relevant to their thoughts; partici-
pants assigned to the “control group” would say as quickly as possible
the color in which the word was printed. The written description ex-
plained, “Responses of people in the control group will provide a base-
line needed to interpret the responses of the people in the test group.
Therefore, it is important that you do your best no matter to which
group you have been assigned.”

Once subjects read this information, the experimenter consulted a
chart and informed each subject that she had (ostensibly randomly)
been assigned to the control group, so her task would be to name the
color (red, blue, green, or brown) in which each word appeared. (Inclu-
dion of the test group in the cover story made it plausible not only that
the reaction time measure be taken but also that some of the words be
relevant to thoughts evoked by the broadcast.) To acquaint subjects with
the reaction time task, the experimenter next showed them slides, each
of a row of Xs in one of the four colors, and had subjects practice the
color-naming procedure until they felt comfortable with it.

10 Seven additional students were excluded from the sample and re-
placed, 1 in each empathy condition because of suspicion, 3 in the low-
empathy condition because they reported being unable to maintain
their assigned listening perspective, and 1 in each empathy condition
because of failure to give a clear response to the request for help.
Manipulation of empathy. Before listening to the "News From the Personal Side" broadcast, which presented Katie Banks's need, subjects were instructed to adopt one of two listening perspectives. The perspective instructions were identical to those used Study 2 to create low- and high-empathy conditions, as was the broadcast about Katie's need. The experimenter remained blind to which listening-perspective instructions the subjects received.

Empathic reaction to learning of Katie's need. At the conclusion of the tape, subjects completed the emotional response questionnaire, which listed 24 adjectives describing different emotional states. Subjects were asked to indicate on 7-point scales (1 = not at all, 7 = extremely) how much they had experienced each emotion while listening to the tape. Included in the list were the same six adjectives used in Studies 1 and 4 to measure empathy.

The experimenter then returned and prepared to administer the reaction time measure, only to discover that she had failed to bring the second slide carousel. While she went to get the carousel, the experimenter left the subject two letters to read, explaining that these two letters, which the professor in charge of the study had asked be given to participants, could be read now to save time. To that end, no response form was provided at this point; it was to be provided later.

Measurement of color-naming latencies for reward-, punishment-, and victim-relevant words, and of desire to help Katie. Once subjects had read both letters, the experimenter returned and administered one of the four counterbalanced Stroop slide sets. The experimenter then gave subjects (a) a brief response form on which to indicate how many hours, if any, they wished to volunteer to help Katie; (b) an envelope in which to seal the response form to assure their anonymity; and (c) a reaction time tape. Included in the list were the same six adjectives used in previous studies to reflect feelings of empathy.

Empathy-induced desire to help Katie. To adjust for a severe positive skew in the number of hours volunteered, subjects' desire to help Katie was coded using the same scale used by Coke et al. (1978): 0 = no help volunteered, 1 = 1 hr, 2 = 2 to 3 hr, 3 = 4 to 5 hr, 4 = more than 5 hr. This scale was much less skewed. Consistent with previous research demonstrating a positive relationship between empathy and helping (Coke et al., 1978; Krebs, 1975; Tbi & Batson, 1982), our subjects volunteered more help in the high-empathy condition (M = 1.50) than in the low (M = 0.50), t(46) = 2.94, p < .005.12 Moreover, as expected, scores on the empathy index were positively correlated with the amount of help volunteered, both overall, r(46) = .46, p < .001, and in the low- and high-empathy conditions, r(22) = .58 and .30, ps = .001 and .07, respectively.

Results and Discussion

Perceptions of Katie's need. In both empathy conditions, subjects reported on the reaction questionnaire that Katie's need was very great (overall M = 8.65 on the 9-point scale), with no reliable differences between conditions, t(46) < 1.0. Apparently, the perspective-taking instructions used to manipulate empathy did not prevent subjects in the low-empathy condition from perceiving Katie to be in considerable need.

Effectiveness of the Empathy Manipulation

To assess the effectiveness of the empathy manipulation in inducing empathic feelings for Katie, we compared scores for subjects in the low- and high-empathy conditions on an index of self-reported empathy. This index was created by averaging responses to the six adjectives on the emotional response questionnaire found in previous research to reflect feelings of empathy: sympathetic, moved, compassionate, tender, warm, and softhearted (Cronbach's alpha = .89). As expected, scores on the 7-point empathy index (1 = not at all, 7 = extremely) were lower for subjects in the low-empathy condition (M = 4.54) than in the high (M = 5.38), t(46) = 2.10, p < .03, one-tailed, indicating that the empathy manipulation successfully induced the intended differences in empathic emotional response to Katie's plight. Still, it should be noted that the level of reported empathy was fairly high even in the low-empathy condition.11

Empathy-induced desire to help Katie. To adjust for a severe positive skew in the number of hours volunteered, subjects' desire to help Katie was coded using the same scale used by Coke et al. (1978): 0 = no help volunteered, 1 = 1 hr, 2 = 2 to 3 hr, 3 = 4 to 5 hr, 4 = more than 5 hr. This scale was much less skewed. Consistent with previous research demonstrating a positive relationship between empathy and helping (Coke et al., 1978; Krebs, 1975; Tbi & Batson, 1982), our subjects volunteered more help in the high-empathy condition (M = 1.50) than in the low (M = 0.50), t(46) = 2.94, p < .005. Moreover, as expected, scores on the empathy index were positively correlated with the amount of help volunteered, both overall, r(46) = .46, p < .001, and in the low- and high-empathy conditions, r(22) = .58 and .30, ps = .001 and .07, respectively.

These results suggested that we had successfully replicated the empathy--helping relationship. Having done this, we turned to the Stroop measure to assess the nature of the goal-relevant cognitions associated with this empathy-induced helping.

Assessing Goal-Relevant Cognitions

To obtain three color-naming latency scores for each subject—one for reward-relevant, one for punishment-relevant, and one for victim-relevant words—we first averaged each subject's color-naming latencies for the four words of each type: reward relevant (M [in milliseconds] = 692.54, SD = 173.61), punishment relevant (M = 709.57, SD = 161.34), victim relevant (M = 685.08, SD = 176.66), and neutral (M = 691.57, SD = 190.70). The counterbalancing within and across slide sets ensured that for each subject all four averages included one response to each of the four colors (red, blue, green, and brown) and that each word--color pairing was equally represented in each empathy condition. To adjust for the individual differences in response latencies that typically occur with reaction time

11 Included on the emotional response questionnaire were eight adjectives that had been found in previous studies to reflect feelings of personal distress caused by witnessing another person suffer: alarmed, grieved, upset, worried, disturbed, perturbed, distressed, and troubled. As in Studies 1 and 3, subjects in this broadcast paradigm seemed to interpret and respond to the distress adjectives more in terms of an other-oriented feeling of distress for Katie than a feeling of being personally distressed by her need.

12 Similarly, a dichotomous measure (0 = no help, 1 = help) revealed that more subjects helped in the high-empathy condition (63; 15 of 24) than in the low (.25; 7 of 24), z = 2.36, p < .01, one-tailed (normal-approximation analysis based on arc sine transformations; see Langer & Abelson, 1972; Winer, 1971).
Turning from the empathy–latency relationships to the latency–helping relationships predicted by the three hypotheses, betas for the three latency measures when regressing helping on these measures, overall and separately for subjects in the low- and high-empathy conditions, are presented in Table 9. As can be seen, the only positive association in the high-empathy condition was the correlation between helping and color-naming latency for the victim-relevant words, which had been predicted by the empathy–altruism hypothesis (β = .62, p < .01). The finding that in the low-empathy condition, in which empathic feelings had not been explicitly aroused, there was not a positive correlation between helping and latency for the victim-relevant words suggested that the positive correlation for victim-relevant words in the high-empathy condition was not due to some general characteristic of these words or their association with helping. The relationship seemed to be empathy specific.

Victim-Relevant Cognitions as Mediators of the Empathy–Helping Relationship

Given the positive association between self-reported empathy and helping and between self-reported empathy and latency for the victim-relevant words, it was possible to examine the degree to which victim-relevant cognitions mediated the relationship between subjects' self-reported empathy and helping. (Because the empathy manipulation did not have a significant overall effect on latency for the victim-relevant cognitions, it was not appropriate to test the role of victim-relevant cognitions in mediating the relationship between the empathy manipulation and helping; see Baron & Kenny, 1986.) Figure 1 provides path coefficients (betas) for the regression equation specified by Baron and Kenny to test mediation. Coefficients are presented for two equations, one for subjects in the low- and one for subjects in the high-empathy condition.

The significant positive path coefficient for the victim-relevant words in the high-empathy condition indicates that victim-relevant cognitions did indeed serve as a mediator of the relationship between self-reported empathy and helping in this condition (β = .51, p < .03). Indeed, in this condition the indirect (mediated) effect was substantial; it accounted for 63% of the total effect of self-reported empathy on helping (Pedhazar, 1982). In the low-empathy condition, however, there was no evidence of an indirect (mediated) effect of self-reported empathy on helping. Instead, there was a strong relationship that was un-

Table 9

<table>
<thead>
<tr>
<th>Type of word</th>
<th>Overalla</th>
<th>Lowb</th>
<th>Highb</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reward relevant</td>
<td>–.20</td>
<td>–.15</td>
<td>–.30</td>
</tr>
<tr>
<td>Punishment relevant</td>
<td>–.20</td>
<td>–.29</td>
<td>–.30</td>
</tr>
<tr>
<td>Victim relevant</td>
<td>.48*</td>
<td>–.06</td>
<td>.62*</td>
</tr>
</tbody>
</table>

* N = 48.  * n = 24.  * p < .01, one-tailed.
Implications of Study 5

Once again, we found no support for either the empathy-specific reward or the empathy-specific punishment hypothesis; there was no evidence of a positive correlation between empathy-induced helping and color-naming latency for either the reward- or punishment-relevant words. Instead, analyses revealed that the increased helping associated with empathy was positively correlated with and mediated by latency to name the color of victim-relevant words. These results provided further support for the empathy-altruism hypothesis, which claims that the goal of empathy-induced helping is to reduce the victim's need. Indeed, given that correlations and tests for mediation are weakened by any error in measurement (Baron & Kenny, 1986)—both our self-report measure of empathy and our color-naming latency measure of salient cognitions almost certainly involved error—the support we found for the empathy-altruism hypothesis seemed striking.

Recall, too, that a positive correlation between amount of helping and color-naming latency for victim-relevant words had also been found in a preliminary study, the interpretation of which was clouded by failure of the empathy manipulation. This failure notwithstanding, the agreement between our results and the results of the preliminary study increased our confidence in the evidence for the empathy-altruism hypothesis found in each study.

At the same time, our analyses also revealed that not all of the observed empathy-helping relationship was mediated by victim-relevant cognitions. Although we have no way of knowing with certainty what accounted for the relatively high levels of self-reported empathy and the relationship between self-reported empathy and helping in the low-empathy condition, previous research (see Batson et al., 1986; Eisenberg & Miller, 1987) revealed that self-presentation concerns (Jones & Pittman, 1982) often account for some, but not all, of women's reported empathy, presumably because the female sex role script (Abelson, 1981; Eagly & Crowley, 1986) says that women should care about others' suffering. Such self-presentation concerns could easily have produced some of the reported empathy in Study 5. At the same time, previous research (Coke et al., 1978; Stotland, 1969) also indicates that such concerns are not likely to account for the difference in reported empathy between the low- and high-empathy conditions. Moreover, Batson et al. (1986) found evidence that only the part of subjects' self-reported empathy that did not reflect self-presentation concerns was associated with motivation to help that appeared altruistic.

Using these previous findings to help us interpret our results, subjects more concerned with presenting themselves as being caring and sympathetic might have also been more likely, when presented with Katie's appeal for help, to comply with her request. In each condition, then, self-presentation might have been at least one source of that part of the empathy-helping relationship that was not mediated by any of our latency measures.

Could Our Results Be an Artifact of Priming?

It may appear that the observed positive correlation between latency for the victim-relevant words and empathy-induced helping could be an artifact of a simple priming effect. After all, the victim-relevant words were by design closely associated with what subjects heard and read about Katie's need. Would not the mere exposure to this material make these words more salient, increasing the color-naming latency? Although such priming certainly might have occurred, there was no evidence that it did. Recall that mean latency for the victim-relevant words was actually slightly lower than mean latency for words of the other three types.

Could priming have occurred specifically in the high-empathy condition, in which the perspective-taking instructions explicitly directed subjects to focus on Katie's feelings? Again, although it might have, there was no evidence that it did. Mean latency for the victim-relevant words was slightly lower in the high-empathy condition ($M = 664.01$) than in the low ($M = 706.16$), $t(46) < 1.0$. A similar lack of effect of the empathy manipulation on mean latency was found for the reward-relevant, punishment-relevant, and neutral words.

Even if there had been priming of the victim-relevant words, either overall or in the high-empathy condition, it is not clear how this priming could account for the observed pattern of correlations predicted by the empathy-altruism hypothesis. Presumably, priming would occur for all subjects (or at least for all subjects in the high-empathy condition). The empathy-altruistic.
ism hypothesis, however, predicted that the color-naming latency for the victim-relevant words would not increase for all subjects, even all subjects in the high-empathy condition. It predicted an increase specifically for those subjects whose increased empathic feelings led them to help. It was this specific, complex pattern of correlations that we observed. These correlations suggested that victim-relevant cognitions were associated not simply with learning of or focusing on Katie’s need but with the link between the empathic emotion evoked by this information and subsequent helping behavior.

Strengths and Weaknesses of the Stroop Procedure

As a research strategy to test the empathy–altruism hypothesis and various egoistic alternatives, the Stroop procedure has both strengths and weaknesses. One strength is that it permitted us to extend research on the empathy–helping relationship into the cognitive domain, testing different motivational hypotheses by exploring possible goal-relevant thoughts associated with empathic emotion. As far as we know, this had not previously been done. Another strength is that the predicted correlations between empathy, helping, and color-naming latencies were subtle and not likely to be produced by experimental demand (Orne, 1962) or self-presentation (Jones & Pittman, 1982). The major weakness of the Stroop procedure is that its validity rests on two plausible but not fully tested assumptions: (a) Color-naming latencies reflect salience of current cognitions and (b) the specific words we used adequately represent the appropriate reward-, punishment-, and victim-relevant cognitions.

Some justification for the first assumption lies in the consistent finding in previous research of increased color-naming latencies specifically for words associated with salient cognitions (Geller & Shaver, 1976; Stroop, 1938; Warren, 1974). Some justification for the second assumption lies in the face validity of the words used and in the finding that they were correctly classified with a high degree accuracy by independent judges. Moreover, there is evidence that color-naming latency for each of the three word sets used is capable of being associated with helping. In Study 5, color-naming latency for both the punishment- and victim-relevant words was, at times, correlated with helping, although the correlation for the punishment-relevant words was the opposite of that predicted by the empathy-specific punishment hypothesis. In a subsequent study by Batson and Dyck (1988) using the same need situation, latency for the reward-relevant words was positively correlated with the helping of subjects scoring relatively high on an intrinsic, end orientation to religion (Allport & Ross, 1967; Batson & Ventis, 1982).

GENERAL CONCLUSIONS

When the results of all five studies are considered, a remarkably consistent pattern emerges. Using different need situations, techniques for operationalizing empathy, and dependent measures, our results consistently conformed to the pattern predicted by the empathy–altruism hypothesis. In no study did the results show a pattern predicted by either the empathy-specific reward or the empathy-specific punishment hypothesis.

Four of the studies were designed to test the empathy-specific punishment hypothesis. Studies 2 through 4 used different techniques to provide justification for not helping, and Study 5 assessed the salience of punishment-relevant cognitions for empathically aroused helpers. We failed to find support for the empathy-specific punishment hypothesis in any study. When these results are considered in conjunction with the results of studies of the empathy–helping relationship using an ease of physical escape manipulation (Batson et al., 1981, 1983; Toi & Batson, 1982) and studies eliminating the possibility of social punishment (Fultz et al., 1986), the claim that the motivation to help evoked by empathy is directed toward the egoistic goal of avoiding empathy-specific punishments seems very doubtful. As with a claim for the existence of unicorns, we cannot categorically say that it is wrong, but we have looked hard in a number of likely places to find supporting evidence and have found none.

Two studies were designed to test the empathy-specific reward hypothesis. Study 1 measured mood after relief of a victim’s need was or was not the result of one’s own helpful act; Study 5 assessed the salience of reward-relevant cognitions for empathically aroused helpers. We found no support for the empathy-specific reward hypothesis in either study. Thus, the claim that the motivation to help evoked by empathy is directed toward the egoistic goal of gaining empathy-specific rewards also seems doubtful.

As we noted when discussing the results of Study 1, however, there is a second version of the empathy-specific reward hypothesis that was not clearly addressed in that study, nor was it clearly addressed in Study 5. Cialdini et al. (1987; see also Schaller & Cialdini, in press) claimed that it is not the needs but the need for the rewards that is empathy specific. Feeling empathy for a person who is suffering involves a state of temporary sadness or depression, which can be relieved by any mood-enhancing experience, including obtaining the social and self-rewards that accompany helping. Cialdini et al. (1987) provided some evidence consistent with this view, but recent research reported by Dovidio, Allen, and Schroeder (1987) and Schroeder, Dovidio, Sibicky, Matthews, and Allen (in press) seems to challenge the Cialdini et al. claim. Still, at this point the possibility that a negative-state relief version of the empathy-specific reward hypothesis can account for the empathy–helping relationship cannot be entirely ruled out.

Moreover, the empathy-specific reward and empathy-specific punishment hypotheses are not the only new egoistic explanations proposed to account for the empathy–helping relationship. Piliavin, Dovidio, Gaertner, and Clark (1981) offered an empathy–cost explanation. They claimed that the empathically aroused individual is motivated to avoid not present but anticipated future aversive empathic arousal produced by the knowledge that the person in need is continuing to suffer. As far as we know, this empathy–cost explanation has yet to be tested directly.

Clearly, before the empathy–altruism hypothesis is accepted, more evidence is needed. At the same time, the consistency and diversity of evidence for the claim that empathic emotion evokes altruistic motivation is becoming hard to ignore. Plausible egoistic explanations of the evidence are increasingly difficult to find. The empathy–altruism hypothesis certainly seems to be the most parsimonious explanation for the results of the five studies reported here. More and more, it appears that the motivation to help evoked by feeling empathy is at least
partly altruistic. If it is, then psychologists will have to make some fundamental changes in their conceptions of human motivation and, indeed, of human nature.

References


Received October 15, 1986
Revision received February 16, 1988
Accepted February 17, 1988

---

**Low Publication Prices for APA Members and Affiliates**

**Keeping You Up-to-Date**
All APA members (Fellows, Members, and Associates) receive — as part of their annual dues — subscriptions to the *American Psychologist, the APA Monitor, and Psychology Today.*

High School Teacher and Student Affiliates receive subscriptions to the *APA Monitor and Psychology Today,* and they can subscribe to the *American Psychologist* at a significantly reduced rate.

In addition, all members and affiliates are eligible for savings of up to 50% on other APA journals, as well as significant discounts on subscriptions from cooperating societies and publishers (e.g., the British Psychological Society, the American Sociological Association, and Human Sciences Press).

**Essential Resources**
APA members and affiliates receive special rates for purchases of APA books, including the *Publication Manual of the APA, the Master Lectures,* and APA's *Guide to Research Support.*

**Other Benefits of Membership**
Membership in APA also provides eligibility for low-cost insurance plans covering life: medical and income protection; hospital indemnity; accident and travel; Keogh retirement; office overhead; and student/school, professional, and liability.

For more information, write to American Psychological Association.
Membership Records, 1200 Seventeenth Street NW, Washington, DC 20036