Social Loafing and Expectancy-Value Theory

James A. Shepperd
Kevin M. Taylor
University of Florida

Expectancy value theory holds that goal-directed behavior is a function of (a) expectations—the belief that performance depends on effort, (b) instrumentality—the belief that outcome depends on performance, and (c) outcome value—the value attached to achieving the outcome. The present research provides a direct test of two factors involved in the instrumentality component: the contingency between individual performance and group performance and the contingency between group performance and group outcome. Experiment 1 revealed that collective participants worked hard when they perceived a contingency between individual performance and group performance. Experiment 2 revealed that collective participants worked hard when they perceived a contingency between group performance and the group outcome. Taken together, the results confirm the importance of high instrumentality in eliminating social loafing.

Many human endeavors cannot be accomplished individually, requiring instead that people combine their efforts toward a common goal. Erecting a building such as the National Cathedral or the Eiffel Tower, playing a team sport such as football or baseball, and completing a group project such as writing a group paper all require that individuals pool their efforts to produce the desired outcome. However, group performance researchers have repeatedly observed that individuals exert less effort when their efforts are pooled than when their efforts are considered individually (e.g., Latané, Williams, & Harkins, 1979). Latané et al. (1979) coined the term "social loafing" to describe the reduction in effort of people working collectively as opposed to coactively and described it as a social disease.

In the last two decades, researchers have documented a variety of causes of social loafing and related phenomena (e.g., free riding) and have proposed several cures (see Karau & Williams, 1993; Shepperd, 1993, for a review). More important, there is growing recognition that low motivation and effort in collective settings is best conceptualized within an expectancy-value theoretical framework (Karau & Williams, 1993; Kerr, 1983, 1986; Shepperd, 1993, 1995; Stroebe & Frey, 1982; Williams & Karau, 1991). The present research examines the instrumentality component of an expectancy-value model of effort motivation in collective settings.

Expectancy-Value Theory

Expectancy-value theories typically consist of three components: expectancy, instrumentality, and value (Karau & Williams, 1993; Mitchell, 1974, 1982; Porter & Lawler, 1968; Vroom, 1964). The expectancy component refers to the perception that performance is contingent on effort (i.e., that greater efforts will result in better performance). For example, a student may believe that if he or she works hard, he or she can write a good term paper (high effort expectancy). Alternatively, he or she may believe that writing a good paper is beyond his or her ability or that the amount of effort that he or she devotes to writing a paper and the quality of the paper that emerges are unrelated (low effort expectancy).

The instrumentality component refers to the perception that the consequence of the performance outcome is contingent on performance (i.e., that performance will determine the outcome). In a sense, it also reflects a type of expectation, but an expectation about performance rather than an expectation about effort. In the term paper example, the student may believe that a good paper will receive a good grade and that a poor paper will receive a bad grade (high instrumentality). On the other

Authors' Note: We thank Kip Williams for insightful comments on an earlier version of this article. We also thank Alicyn Baker, Amy Ryan, Jennifer Sager, and Holly Siegel for serving as experimenters and Jeff Smith for helpful comments during Experiment 1. Correspondence concerning this article should be sent to James A. Shepperd, Department of Psychology, P.O. Box 112250, University of Florida, Gainesville, FL 32611-2250; e-mail: shepperd@psych.ufl.edu.

PSPB, Vol. 25 No. 9, September 1999 1147-1158

hand, the student may view the teacher as capricious and thus perceive no relationship between the quality of the paper and the grade received (low instrumentality).

The value component refers to how much value or importance the person attaches to achieving the outcome of the performance minus any costs involved. In the term paper example, the student's course grade may depend largely on the grade received on the term paper, thus making the term paper grade important (high value). Alternatively, the term paper grade may have a negligible impact on the student's course grade (low value). Of importance, value does not depend solely on how important or rewarding the person regards the outcome; it also depends on the costs (psychological and material) associated with achieving the outcome (Shepperd, 1993). Thus, outcome value represents the difference between the reward for achieving the outcome and the cost of achieving the outcome. For many events, there are multiple values. There may be value attached to completing a task, such as completing the term paper, independent of any external rewards received for performance.

When viewed together, effort motivation can be expressed as the product of expectation, instrumentality, and outcome value. Effort motivation in collective settings should be high when people (a) perceive a contingency between their effort and the performance, (b) perceive a contingency between performance and the outcome, and (c) value the outcome (the benefits associated with contributing or achieving the outcome exceed the costs of contributing). Effort motivation reflects how much effort a person is willing to exert on a task or toward a goal.

Of importance, effort motivation is a product of the perceived expectancy, instrumentality, and outcome value (Karau & Williams, 1993). The perceptions may only loosely correspond to actual contingencies and value. Moreover, as noted earlier, a given situation may have multiple values and costs each with its own set of expectations and utilities. As a result, the model can lead to actions that are rational to the actor yet appear irrational to observers. For example, a woman may spend an inordinate amount of time working on a task that she realizes she has no hope of successfully completing (the project's effort expectancy is low) or for which any external reward is negligible (outcome value appears low). However, she may devote effort to the task out of a sense of obligation or duty, because performing the task is intrinsically enjoyable, or because important others will notice and supply nontangible rewards (respect, praise, etc.) for high effort. These outcomes, although valued by the person, may go unnoticed by observers, leading to the conclusion that the person's behavior is irrational.

Instrumentality in Collective Settings

In collective settings, where efforts are pooled, the expectancy and value component of expectancy-value models remain the same. However, as noted by Karau and Williams (1993), instrumentality becomes a function of the perceived contingency between (a) individual performance and group performance, (b) group performance and group outcome, and (c) the group outcome and the individual outcome. Our interest in this research was to examine the first and second of these contingencies.

Several studies have examined the contingency between individual performance and group performance by, for example, manipulating the group size or the number of contributors (Ingham, Levinger, Graves, & Peckham, 1974; Latané et al., 1979; Petty, Harkins, Williams, & Latané, 1977; Weldon & Mustari, 1988; Williams, Harkins, & Latané, 1981) or the dispensability of personal efforts (Kerr, 1983; Kerr & Bruun, 1983; Williams & Karau, 1991). The typical finding is that people exert greater effort when they perceive their individual efforts as indispensable yet loaf when they perceive their efforts as dispensable. For example, in a study by Kerr and Bruun (1983), participants working in pairs believed that they were more or less able than their partner at an air-blowing task. In addition, participants believed that the pair's performance would be equal to the performance of the least able member of the pair (a conjunctive task) or would equal the performance of the most able member of the pair (a disjunctive task). Kerr and Bruun found that high-ability participants exerted greater effort on the disjunctive task than on the conjunctive task, whereas low-ability participants exerted greater effort on the conjunctive task than on the disjunctive task. In short, when participants saw their contributions as crucial (indispensable) for a good collective performance, they worked harder. When they saw their contributions as dispensable, they loafed.

Few studies have examined the effect of the contingency between group performance and group outcome on effort in collective settings. Within organizational psychology, some studies have manipulated the effect of group performance on group outcome. However, these studies have most often examined self-reported rather than actual effort (e.g., Mitchell, 1982). In the few studies that have examined actual effort, instrumentality was naturally occurring and thus uncontrolled, or individual efforts were easily identifiable and thus open to evaluation by the experimenter (e.g., Ilgen, Nebeker, & Pritchard, 1981; Peters, 1977; Pritchard & De Leo, 1973).

The Present Research

Expectancy-value theory provides a simple, elegant framework for understanding and predicting low moti-

vation and effort in collective settings. A growing number of researchers are incorporating the notion of expectancies in their investigations of effort in collective settings (e.g., Karau & Williams, 1993; Kerr, 1983, 1986; Mullen & Baumeister, 1987; Sanna, 1992; Sanna & Pusecker, 1994; Sanna & Shotland, 1990; Shepperd, 1993, 1995; Stroebe & Frey, 1982; Williams & Karau, 1991). Moreover, virtually every study within the social loafing literature appears explainable, within an expectancy-value framework. That is, every study appears to have examined effort or productivity of people working individually or collectively by manipulating one or more of the three components of expectancy-value theory.

Nevertheless, much research on social loafing was conducted prior to the general recognition of the utility of expectancy-value theory for predicting social loafing. Only a few studies have examined the effect of instrumentality on collective effort, and they typically have done so indirectly by manipulating the dispensability of individual contributions. What follows are two studies that provide a direct test of instrumentality on effort. Experiment 1 manipulated the contingency between individual performance and group performance. Experiment 2 manipulated the contingency between group performance and group outcome. Both experiments included a manipulation of experimenter evaluation, thereby permitting a comparison of the findings with prior social loafing research.

EXPERIMENT 1

Experiment 1 examined the effect of instrumentality by exploring how the link between individual performance and group performance affect effort using a variation on a paradigm employed frequently in previous research (Harkins & Szymanski, 1988). Participants generated uses for an object while working as part of a collective. In addition, participants were told that a good collective performance would be rewarded, whereas other participants were given no such information. To manipulate instrumentality, participants in the reward condition learned through a series of practice trials that the probability that their group could achieve a good collective performance was either high or low. To permit comparisons with past research, we also manipulated whether individual contributions could be evaluated. Specifically, participants believed that their individual uses would or would not be counted by the experimenter at the end of the experiment. The design was thus a 2 (evaluation vs. no-evaluation) $\times$ 3 (high instrumentality vs. low instrumentality vs. control) between-participants factorial. The two control cells contained no instrumentality manipulation. The no-evaluation / control cell served as the comparison cell for evaluating the hypotheses.

We hypothesized that the prospect of experimenter evaluation would lead to high effort regardless of the instrumentality manipulation. Again, our rationale was that experimenter evaluation would serve to make the uses-generation task an individual one rather than a collective one, leading participants to work hard to produce a positive experimenter evaluation (or avoid a negative evaluation) regardless of whether an external reward was provided for contributing (see Shepperd, 1993). Thus, we predicted that participants in each of the three evaluation conditions would generate more uses than would participants in the no-evaluation / control condition. We also hypothesized that participants would work hard when contributions could not be evaluated, provided they believed that a good collective performance was possible and dependent on their efforts. Thus, we predicted that participants in the no-evaluation / high-instrumentality condition would work hard, generating more uses than participants in the no-evaluation / control condition. By contrast, we hypothesized that participants would not work hard if they believed that a good collective performance was unlikely regardless of how many uses they individually generated. Stated otherwise, we hypothesized that when participants perceived no contingency between their individual performance and achieving the reward, participants would not work hard. Thus, we predicted that participants in the no-evaluation / low-instrumentality condition would not differ from participants in the no-evaluation / control condition in the number of uses generated.

METHOD

Participants

Participants were 103 undergraduates (35 male, 68 female) randomly assigned to conditions. When possible, participants were run in groups of five. However, in some sessions, only three or four participants arrived on time, resulting in smaller session sizes. All participants participated to fulfill an obligation in introductory psychology.

Procedure

A female experimenter met participants in a waiting room and explained that the study involved performing a brainstorming task and that participants were to generate uses for an object. Participants were told that the experiment was concerned with the quantity and not the quality of ideas and that participants should generate as many uses as possible. In addition, participants learned that they would be performing the brainstorming task as part of a group consisting of the other participants in

their session. When a session contained less than three participants, the participants were told that they were still generating uses as part of a group and that their uses would be combined with those of other participants to fill out their group. After receiving all instructions, participants were escorted to enclosed cubicles, making it impossible for participants to monitor one another. Participants then received a stack of 2-inch × 3-inch blank slips of paper with instructions to write one use per slip.

Evaluation manipulation. Similar to previous research, participants learned that their individual efforts would or would not be evaluated. In the evaluation condition, the experimenter informed participants that she would count the number of uses that participants generated before placing their slips in a common box with the slips generated by other members of their group. Participants in the no-evaluation condition received instructions to place their uses through the slit of the common box with the slips from other members of their group. The experimenter also stressed to no-evaluation participants that she had no way of determining how many uses each participant individually generated. In truth, the slips of paper in each cubicle differed in size, making it possible to distinguish each participant's individual contributions.

Instrumentality manipulation. Crossed with the evaluation manipulation was the manipulation of instrumentality. Participants in the high- and low-instrumentality conditions were told that if their group performed well on the brainstorming task, members of their group would receive a prize. Specifically, the experimenter explained to high- and low-instrumentality participants that their group was one of 40 participating in the study and that if the number of uses generated by their group on the final brainstorming task was among the top 10% of uses generated by all groups, they would receive a prize, either certificates for free pizza or free movie passes. Participants in the control condition learned nothing about the possibility of winning a prize for performing well on the brainstorming task.

Participants next performed the four practice trials in which they generated uses for a bed sheet, comb, toothbrush, and shoebox. Prior to each practice trial, the experimenter reminded participants to generate as many uses as they could during each trial. After each practice trial, the experimenter combined the uses generated by each participant and took them to an adjoining room where they were counted by an assistant. In the evaluation condition, the experimenter counted the slips before combining them. In the no-evaluation condition, participants placed their slips through the slit in the common box while the experimenter looked away, thus making the individual contributions apparently anonymous.

Shortly after each trial, participants in the high- and low-instrumentality conditions received written performance feedback privately from the assistant. The feedback was ostensibly presented as a means for participants to monitor their group's performance and to give them some idea of the likelihood that their group would perform well enough on the final trial to receive the prize. In truth, the feedback was false. In the high-instrumentality condition, the feedback was designed to create the expectation within participants that their group was close to generating enough uses to be in the top 10% of groups and that high effort on their part was crucial. In the low-instrumentality condition, the feedback was designed to create the expectation within participants that their group was unlikely to generate enough uses to be in the 10%, even if they individually worked very hard.

High-instrumentality participants learned that their group's performance placed them in the 90th percentile on the first trial, in the 82nd percentile on the second trial, in the 92nd percentile on the third trial, and the 84th percentile on the fourth trial. The implication was that high-instrumentality participants were performing well and that if they continued to work hard, there was a good chance that their group would perform well enough on the final trial to win the prize. Low-instrumentality participants learned that their group's performance placed them in the 86th percentile on the first trial, in the 68th percentile on the second trial, in the 66th percentile on the third trial, and the 49th percentile on the fourth trial. The implication was that low-instrumentality participants were performing modestly and that it was unlikely that their group would perform well enough on the final trial to win the prize. Control participants also received four practice trials and were reminded prior to each trial to generate as many uses as they could. However, they received no feedback regarding how their group performed on the practice trials.

After receiving feedback regarding the fourth practice trial, high- and low-instrumentality participants (but not control participants) completed a brief questionnaire asking (a) "How confident are you that your group will perform well on the brainstorming task?" (b) "How necessary is your effort to achieving a good group performance?" and (c) "If the task were an individual one instead of a group one, and you were working alone, how likely is it that you could generate enough uses to win the reward?"

Next, the experimenter announced the beginning of the final trial. Participants were instructed to generate as many uses as they could for a knife during the 15-minute trial. At the end of the trial, the experimenter instructed

participants to stop and collected the slips of paper. All participants then completed a final questionnaire consisting of distractor and manipulation check items and were debriefed. Finally, the names of all participants (including those in the control condition) were entered into a lottery to receive a prize; either a certificate for a free pizza or free movie passes. No participant voiced any suspicion of the manipulations or procedures during the debriefing.

RESULTS AND DISCUSSION

Preliminary analyses revealed no difference between males (M = 32.8) and females (M = 24.6) in the number of uses generated during the brainstorming task, $F(1, 91) = 2.40$, $p > .12$. Other analyses revealed that the number of participants run per session (M = 4.3, SD = .79) was uncorrelated with number of uses generated by participants, $r(101) = –.07$. Therefore, sex and session size were excluded as variables. Unless otherwise indicated, subsequent analyses were conducted using a 2 (evaluation vs. no-evaluation) × 3 (high instrumentality vs. low instrumentality vs. control) between-participants analyses of variance (ANOVA).

Manipulation Check Items

Two items assessed the effectiveness of the evaluation manipulation. Both yielded a single, significant effect in the predicted direction. Participants in the evaluation condition (M = 5.6, SD = 2.6) agreed more than participants in the no-evaluation condition (M = 3.7, SD = 2.5) that the experimenter would know or would be able to tell how well they individually performed on the brainstorming task, $F(1, 102) = 13.74$, $p < .0003$. Similarly, participants in the evaluation condition (M = 5.8, SD = 2.8) agreed more than participants in the no-evaluation condition (M = 3.6, SD = 2.3) that the experimenter knew or would know how many uses they personally generated, $F(1, 102) = 20.22$, $p < .0001$.

Two items administered only in the high- and low-instrumentality conditions revealed that the instrumentality manipulation also was quite effective. Specifically, participants were more confident in the high-instrumentality condition (M = 6.7, SD = 1.2) than in the low-instrumentality condition (M = 5.3, SD = 1.9) that their group would perform well on the brainstorming task, $F(1, 64) = 11.78$, $p < .002$. Likewise, participants in the high-instrumentality condition (M = 6.2, SD = 1.5) were more likely than participants in the low-instrumentality condition (M = 4.2, SD = 2.1) to believe that their group would win the prize, $F(1, 64) = 19.61$, $p < .0001$.

Results from a third item also illustrate the effectiveness of the instrumentality manipulation. Specifically, participants indicated whether they believed that their group's performance on the final trial would be in the top 10% of all groups participating in the study. Analysis yielded a single significant effect of instrumentality, $F(2, 101) = 11.61$, $p < .0001$. Participants in the high-instrumentality condition (M = 6.4, SD = 1.3) agreed more than participants in the control condition (M = 5.3, SD = 1.3) that their group would be among the top 10%, $t(1, 66) = 2.77$, $p < .05$. Participants in the control condition, in turn, agreed more than participants in the low-instrumentality condition (M = 4.6, SD = 1.8) that their group would be among the top 10%, $t(1, 67) = 2.05$, $p < .02$.

Although the instrumentality manipulation affected participants' perceptions of how well their group would do on the brainstorming task, it did not affect participants' perceptions of how well they individually would do on the brainstorming task. That is, there were no differences across conditions in responses to an item asking participants the extent to which their individual efforts were necessary to achieve a good collective performance (all ps > .25; grand M = 7.0, SD = 2.4). More specifically, low-instrumentality participants (M = 7.2, SD = 2.6) were just as likely as high-instrumentality participants (M = 6.7, SD = 2.1) to regard their efforts as necessary for a good collective performance. In addition, there was no difference across conditions in responses to an item asking participants to rate the likelihood that they could generate enough uses to win the prize if winning was based on individual performance rather than group performance (all ps > .50; grand M = 5.9, SD = 2.1). Specifically, low-instrumentality participants (M = 6.0, SD = 2.3) were just as likely as high-instrumentality participants (M = 5.9, SD = 1.8) to believe that they could generate enough uses to receive the prize if the brainstorming task were based on individual performance rather than on the group's performance. In short, the instrumentality manipulation clearly affected participants' perceptions of their group's ability yet did not affect participants' perceptions of their own ability.

We included several additional items that examined perceptions during the task. Analyses revealed no difference across conditions in participants' reports of how motivated they were (grand M = 5.9, SD = 2.0), how well they believed they performed on the final brainstorming task (grand M = 5.6, SD = 1.6), or how much effort they reported exerting (grand M = 6.4, SD = 2.0). There was, however, a significant effect of instrumentality for an item asking participants how difficult they found the task, $F(1, 97) = 4.62$, $p < .01$. Post hoc analyses using the Student Newman Keuls multiple range test revealed that control participants (M = 4.5, SD = 2.3) found the final brainstorming task more difficult than did high-instrumentality (M = 3.4, SD = 2.0) and low-instrumentality (M = 3.0, SD = 1.9) participants but that

high- and low-instrumentality participants did not differ in their judgment of the difficulty of the final task. Although unanticipated, this finding raises no interpretation problems for the primary dependent measure.

In sum, both the evaluation and instrumentality manipulations were effective. Moreover, although control participants found the final brainstorming trial more difficult than did high- and low-instrumentality participants, there were no differences across conditions for the motivation to perform well, effort exerted during the final trial, or judgments of performance during the final trial. We now turn to the primary dependent measure.

Number of Uses Generated

The number of uses generated in each condition during the final trial are presented in Figure 1. Specific hypotheses were examined using a series of orthogonal contrasts. The first hypothesis was that participants in the control condition would generate more uses when their efforts could be evaluated than when their efforts could not be evaluated. Consistent with this prediction, evaluation / control participants (M = 28.4, SD = 15.5) generated significantly more uses than did no-evaluation / control participants (M = 20.0, SD = 9.7), $t(1, 97) = 1.94$, $p < .05$, $\eta^2 = .04$, one-tailed. This finding replicates the standard social loafing effect.

We also predicted that in the low-instrumentality condition, participants would generate more uses when efforts could be evaluated than when efforts could not be evaluated but that in the high-instrumentality condition, evaluation and no-evaluation participants would not differ in the number of uses generated. Consistent with predictions, participants in the low-instrumentality condition generated more uses when their efforts could be evaluated (M = 37.9, SD = 17.3) than when they could not be evaluated (M = 21.11, SD = 11.7), $t(1, 97) = 3.82$, $p < .01$, $\eta^2 = .13$. Also consistent with predictions, participants in the high-instrumentality / evaluation condition (M = 28.4, SD = 12.3) and high-instrumentality / no-evaluation condition (M = 28.4, SD = 7.3) did not differ in the number of uses generated, $t(1, 97) = .00$, $p = .99$, $\eta^2 = .00$. Finally, as expected, no-evaluation / high-instrumentality participants generated more uses than did no-evaluation / control participants, $t(1, 97) = 1.88$, $p < .05$, $\eta^2 = .04$, one tailed, whereas no-evaluation / low-instrumentality participants did not differ from no-evaluation / control participants in the number of uses generated, $t(1, 97) = .24$, $p = .80$, $\eta^2 = .00$.

We recognized the possibility that our instrumentality manipulation may have unintentionally affected participants' perceptions of the desirability of the reward. Specifically, it is possible that participants may have rated the reward as less valuable in the low-instrumentality condition to cope with the fact that they were unlikely to receive the reward. To examine this possibility, participants in the high- and low-instrumentality conditions completed an additional item asking how much they would like to win the pizza (1 = *not much*, 9 = *very much*). Analysis of the item revealed no main effects or interaction of the evaluation and instrumentality manipulations (all ps > .20). That is, participants did not differ across instrumentality conditions in how much they wanted to receive the prize.

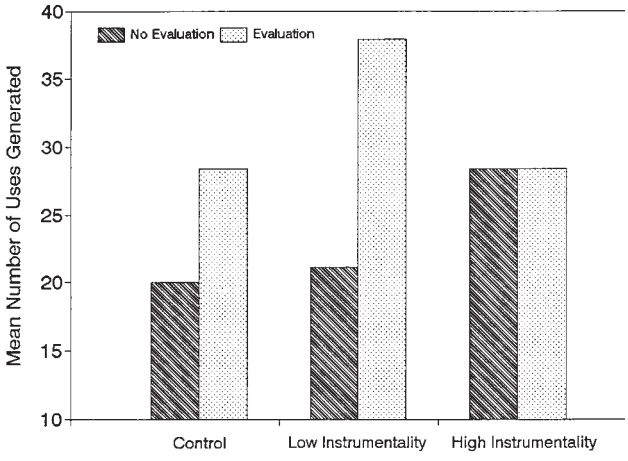

Figure 1 Number of uses generated as a function of instrumentality and evaluation.

In sum, consistent with prior research and theory on social loafing (see Karau & Williams, 1993; Shepperd, 1993), participants exerted high effort on the final brainstorming task when their individual efforts would be evaluated by the experimenter, irrespective of whether they anticipated a prize for a good group performance (i.e., the control condition) and irrespective of their expectations regarding the likelihood that their group would generate enough uses to receive the prize (i.e., the high- and low-instrumentality conditions). In the no-evaluation condition, by contrast, participants exerted high effort only when they believed it likely that their group would generate enough uses to receive the prize (i.e., in the high-instrumentality condition). When participants believed that their group was unlikely to generate enough uses to receive the prize (the low-instrumentality condition) or when there was no mention of a prize (the control condition), participants loafed.

Within the evaluation condition, the number of uses generated by low-instrumentality participants was strikingly high. Moreover, a similar pattern emerged when we pilot tested the procedures. Although post hoc analyses using the Student Newman Keuls multiple range test revealed that these participants did not generate signifi-

cantly more uses than did high-instrumentality and control participants in the evaluation condition, the high output of low-instrumentality/evaluation participants merits attention.[1] Their high effort does not appear due to a feeling of personal responsibility for their group's poor practice trial performance and a desire to demonstrate to the experimenter that they could do better. After all, low-instrumentality participants were just as likely as high-instrumentality participants to believe that they personally could achieve the prize if the brainstorming task were an individual one rather than a group one, implying that they viewed the poor group performance as caused by other group members. Nor does the greater effort seem to result from a belief by low-instrumentality/evaluation participants that a poor group performance on the final task, although caused by others, would be erroneously attributed to them by the experimenter. After all, participants in the evaluation conditions knew that the experimenter could monitor who was and who was not loafing, thereby minimizing any personal attributions for a poor collective performance. Moreover, if this explanation were true, then participants in the low-instrumentality/no-evaluation condition would have worked hard too; they did not.

The most likely explanation is that an attempt to compensate for their underperforming coparticipants during the practice trials trapped these low-instrumentality participants into working hard on the final trial to appear consistent. Accordingly, these low-instrumentality participants worked hard on the initial trial because they could be evaluated. The feedback of a mediocre collective performance led them to increase their effort during the second, third, and fourth trials. After four trials of working hard, however, participants may have reasoned that they had established a high performance expectation in the eyes of the experimenter. They consequently had to work extremely hard on the final task because anything less would be perceived as inconsistent (and as loafing) in the eyes of the experimenter.

EXPERIMENT 2

Experiment 1 demonstrated that the contingency between individual performance and group performance influences the effort exerted on a task. Experiment 2 examined a second factor of instrumentality (i.e., whether the contingency between group performance and group outcome exercises a similar influence on individual effort). Similar to Experiment 1, participants generated uses for an object as part of a collective and believed that the uses that they individually generated would or would not be counted by the experimenter. In addition, some collectives were offered a reward to work hard, whereas others were not. Within the reward condition, participants learned either that there was a high or low probability that collectives that performed at a high level would be rewarded. Thus, as in Experiment 1, the design was a 2 (evaluation vs. no-evaluation) × 3 (high instrumentality vs. low instrumentality vs. control) between-participants factorial.

Consistent with prior research, we hypothesized that the prospect of experimenter evaluation would lead to high effort regardless of the instrumentality manipulation. Our rationale was that experimenter evaluation would serve to make the uses-generation task an individual one rather than a collective one, leading participants to work hard to produce a positive experimenter evaluation (or avoid a negative evaluation) regardless of whether an external reward was provided for contributing (see Shepperd, 1993). We also hypothesized that participants would work hard even when contributions could not be evaluated, provided there was a reasonable chance of obtaining a desired reward. Thus, we predicted that participants in the high-instrumentality condition would work hard in both the evaluation and no-evaluation conditions. By contrast, we predicted that participants would not work hard if the prospect of obtaining a desired reward was remote. Thus, we predicted that participants in the evaluation/low-instrumentality condition would work hard, whereas participants in the no-evaluation/low-instrumentality condition would not.

METHOD

Participants

Participants were 90 undergraduate students (38 male, 52 female) randomly assigned to conditions. When possible, participants were run in groups of three. However, in some sessions, fewer than 3 participants arrived on time for the experiment, resulting in only 1 or 2 participants run in the session. All participants participated to fulfill an obligation in introductory psychology.

Procedure

The procedures were generally similar to those of Experiment 1. Participants learned that they would perform a brainstorming task and that each participant in the session was part of the same group. Unlike Experiment 1, there were no practice trials.

Evaluation manipulation. Similar to Experiment 1, the experimenter told participants that their individual efforts would or would not be evaluated. The evaluation manipulation in Experiment 2 was identical to Experiment 1 with the exception that the slips of paper in the no-evaluation condition differed in color rather than size. A lid on the box with a narrow slit prevented no-evaluation participants from knowing that their slips of

paper differed in color from the slips of other participants.

Instrumentality manipulation. Fully crossed with the evaluation manipulation was an instrumentality manipulation. Participants in the high-instrumentality condition were told that 10 groups were participating in this experiment and that if their group was among the 7 groups that generated the most uses, each member of the group would receive a certificate for a free slice of pizza. Thus, other things being equal, they had a 70% chance of receiving the pizza. Participants in the low-instrumentality condition were told that at least 40 groups were participating in the experiment. If their group was among the 4 groups that generated the most uses, their names would be entered into a lottery. At the end of the experiment, one name would be selected from the lottery to receive a certificate for a free slice of pizza. The experiment further explained that, other things being equal, there was only a 1 in 200 chance of winning pizza at the moment but if their group was among the top 4 groups, there was a 1 in 20 chance that their name would be drawn to win pizza.[2] Participants in the control condition were told nothing of the possibility of winning pizza.

When the experimenter was sure that participants understood all instructions, she announced that they would have 10 minutes to generate uses for a knife. She then reminded participants to generate as many uses as possible during the time period. When 10 minutes had elapsed, the experimenter collected the slips of paper. In the no-evaluation condition, the experimenter went to each cubicle and directed participants to place their slips in a group box while she appeared inattentive. In the evaluation condition, the experimenter counted and recorded the number of slips prior to placing them in the box. All participants then completed a brief questionnaire that included manipulation check items and items measuring participants' perceptions of the brainstorming task. After completing the questionnaire, participants were debriefed. The names and phone numbers of all participants (including control participants) were collected and entered into a lottery for the pizza. No participant voiced any suspicion of the manipulations or procedures during the debriefing.

RESULTS AND DISCUSSION

Preliminary analyses revealed no reliable main effects or interactions involving sex for the primary dependent measure. Males (M = 20.6) and females (M = 21.1) did not differ in the number of uses they generated. Likewise, the number of participants run per session (M = 2.5, SD = .67) was uncorrelated with the number of uses generated, $r(88) = -.09$. Thus, sex and session size were excluded as variables. Unless otherwise indicated, data were analyzed using a 2 (evaluation vs. no-evaluation) × 3 (high instrumentality vs. low instrumentality vs. control) between-participant analysis of variance (ANOVA). In addition, with one exception noted below, responses to all manipulation check items were made using 9-step scales typically anchored by 1 = *strongly disagree* and 9 = *strongly agree*. Finally, variations in the cell size for some analyses are due to one participant providing incomplete data.

Manipulation Check Items

Two items assessed the evaluation manipulation and both yielded a significant effect in the predicted direction. Participants in the evaluation condition (M = 6.1, SD = 2.3) agreed more than participants in the no-evaluation condition (M = 4.2, SD = 2.2) that the experimenter knew how well they individually performed, $F(1, 84) = 15.92$, $p < .001$. Participants in the evaluation condition (M = 7.1, SD = 2.4) also agreed more than participants in the no-evaluation condition (M = 3.8, SD = 2.5) that the experimenter knew how many uses they personally generated, $F(1, 84) = 42.44$, $p < .001$.

Two items also assessed the effectiveness of the instrumentality manipulation. The first item asked participants to estimate the probability (from 0% to 100%) that they would win the pizza if they personally worked hard. The second item asked participants how likely they were to win the free slice of pizza. Because these items focused specifically on receiving the pizza, they were administered to high- and low-instrumentality participants only. A 2 (evaluation vs. no-evaluation) × 2 (high vs. low instrumentality) between-participants ANOVA revealed significant main effects for both items. Analysis of the probability item yielded a significant main effect of instrumentality, $F(1, 55) = 80.31$, $p < .001$. Participants in the high-instrumentality condition (M = 71.5%, SD = 13.8) estimated a higher probability of winning than did participants in the low-instrumentality condition (M = 24.1%, SD = 25.2). Analysis of the second item yielded a similar main effect, $F(1, 55) = 76.70$, $p < .001$. Participants in the high-instrumentality condition (M = 6.0, SD = 1.7) agreed more than participants in the low-instrumentality condition (M = 2.4, SD = 1.5) that they would win the pizza. Of importance, analysis of the second item also yielded a significant Evaluation × Instrumentality interaction, $F(1, 55) = 5.51$, $p < .03$. Post hoc analyses revealed that the interaction was attributable to a greater difference in estimated likelihood of winning among high- (M = 6.5, SD = 2.0) and low- (M = 1.9, SD = .89) instrumentality participants in the no-evaluation condition, $t(28) = 8.01$, $p < .001$, than among high- (M = 5.5, SD = 1.2) and low- (M = 2.9, SD = 2.0) instrumentality participants in the evaluation condition, $t(28) = 4.45$, $p <$

.001. Although this interaction was unexpected, it does not qualify the predicted main effect.

We included several additional items that examined perceptions during the task. Analyses revealed no differences across conditions in agreement to items measuring how motivated participants were (grand $M = 6.2$, $SD = 2.0$), how difficult they found the task (grand $M = 3.6$, $SD = 2.0$), and how much effort they reported exerting (grand $M = 6.0$, $SD = 2.0$).

In sum, the manipulations were clearly effective. We now turn to the primary dependent measure.

Number of Uses Generated

The number of uses that participants generated in each condition are presented in Figure 2. Specific hypotheses were examined using the pooled error term and pooled degrees of freedom in a series of orthogonal contrasts. The first hypothesis was that participants in the control condition would generate more uses when their efforts could be evaluated than when their efforts could not be evaluated. Consistent with this prediction, evaluation / control participants ($M = 20.9$, $SD = 7.0$) tended to generate more uses than no-evaluation / control participants ($M = 16.1$, $SD = 6.6$), $t(1, 84) = 1.63$, $p < .06$, $\eta^2 = .03$, one-tailed. Although only marginally significant, this finding is consistent with Experiment 1 and replicates the standard social loafing effect.

We also predicted that in the low-instrumentality condition, participants would generate more uses when efforts could be evaluated than when efforts could not be evaluated but that in the high-instrumentality condition, evaluation and no-evaluation participants would not differ in the number of uses generated. Consistent with predictions, participants in the low-instrumentality condition generated more uses when their efforts could be evaluated ($M = 24.2$, $SD = 9.3$) than when they could not be evaluated ($M = 16.3$, $SD = 6.9$), $t(1, 84) = 2.67$, $p < .01$, $\eta^2 = .11$. Also consistent with predictions, participants in the high-instrumentality / evaluation condition ($M = 24.4$, $SD = 10.3$) and high-instrumentality / no-evaluation condition ($M = 23.6$, $SD = 7.5$) did not differ in the number of uses generated, $t(1, 84) = .26$, $p = .79$, $\eta^2 = .00$. Finally, as expected, no-evaluation / high-instrumentality participants generated more uses than did no-evaluation / control participants, $t(1, 84) = 2.56$, $p < .01$, $\eta^2 = .07$, whereas no-evaluation / low-instrumentality participants did not differ from no-evaluation / control participants in the number of uses generated, $t(1, 84) = .10$, $p = .93$, $\eta^2 = .00$.

Ancillary Findings

Similar to Experiment 1, participants in the high- and low-instrumentality condition completed an additional item on the final questionnaire asking how much they would like to win the prize (1 = *not much*, 9 = *very much*). We were concerned that the manipulation of instrumentality may have unintentionally affected participants' perceptions of the reward, leading low-instrumentality participants to attach less value to the prize because they viewed the prize as unattainable. Analysis of this item yielded both a main effect of evaluation, $F(1, 55) = 4.13$, $p < .05$, and a main effect of instrumentality, $F(1, 55) = 10.51$, $p < .01$. Both main effects, however, were qualified by a significant Evaluation × Instrumentality interaction, $F(1, 55) = 20.12$, $p < .001$. Post hoc comparisons revealed that in the evaluation condition, high-instrumentality ($M = 5.5$, $SD = 2.2$) and low-instrumentality ($M = 6.1$, $SD = 2.6$) participants did not differ in how much they reported wanting to win the pizza ($t < 1$). By contrast, in the no-evaluation condition, high-instrumentality participants ($M = 6.8$, $SD = 1.9$) were more likely than low-instrumentality participants ($M = 2.6$, $SD = 1.6$) to report that they wanted to win the pizza, $t(28) = 5.51$, $p < .01$.

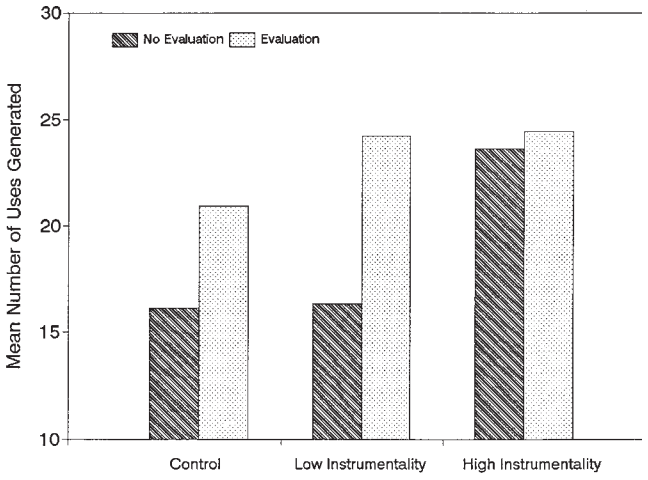

Figure 2 Number of uses generated as a function of instrumentality and evaluation.

At first glance, this interaction seems troubling because it suggests that the manipulation of instrumentality may have inadvertently affected the value that participants placed on performing the task. However, participants reported a low desire to win the pizza only in the no-evaluation condition. In the evaluation condition, participants in the low-instrumentality condition wanted to win the pizza just as much as participants in the high-instrumentality condition, even though winning the pizza, by their own admission, was far less likely for low-instrumentality participants than for high-instrumentality participants. Moreover, it is possible that these reports reflect an attempt to manage impressions and were not participants' true sentiments.

Nevertheless, the fact that the pattern of means for this item corresponded to the pattern of means for the number of uses generated in the high- and low-instrumentality conditions led us to examine more closely whether participants' efforts on the brainstorming task were linked to the value they placed on the reward. We thus reanalyzed the number of uses that participants generated in the four instrumentality conditions after including participants' reports of how much they wanted to win the pizza as a covariate. The covariance analysis revealed that the value that participants placed on winning only modestly predicted the number of uses that participants generated, $F(1, 54) = 3.10$, $p < .09$. More important, when we examined the means adjusted for the covariate, we found that participants in the high-instrumentality condition worked hard regardless of whether their efforts could ($M = 24.6$) or could not ($M = 25.1$) be evaluated, $t < 1$. We also found that, after adjusting for the covariate, participants in the low-instrumentality condition worked hard when their contributions could be evaluated ($M = 25.7$) yet loafed when their contributions could not be evaluated ($M = 13.8$), $t(1, 54) = 3.21$, $p < .01$. Collectively, these findings suggest that although participants differed across conditions in the value they placed on the reward, this did not influence the effect of the instrumentality manipulation on how many uses the participants generated for a knife.

In sum, our instrumentality manipulation (i.e., the likelihood that a good collective performance would be rewarded) affected how much effort participants expended. In the absence of external evaluation, participants worked hard only when instrumentality was high (i.e., when a reward was offered for a good collective performance and the probability of achieving the reward was high).

GENERAL DISCUSSION

The goal of this research was to provide a more direct test of the instrumentality component of expectancy-value theory in explaining low motivation and effort in collective settings. Experiment 1 revealed that no-evaluation participants worked hard only when they perceived a contingency between individual performance and group performance. Experiment 2 revealed that no-evaluation participants worked hard only when they perceived a contingency between group performance and the group outcome. Taken together, these results reveal that people will work hard for a valued outcome if instrumentality is high.

Participants in both experiments worked hard in the evaluation condition regardless of whether instrumentality was high or low. One might ask why the manipulations of instrumentality did not affect individual performance when performances could be evaluated. We propose that the opportunity for experimenter evaluation transformed the task from a collective one to an individual one, shifting the expectations, instrumentality, and value toward achieving a favorable experimenter evaluation rather than realizing achieving the reward. As such, participants worked hard even when there was no added reward (the control condition) or the reward was remote (the low-instrumentality condition).

An intriguing finding in Experiment 1 was the enormous effort exerted by participants in the evaluation/low-instrumentality condition. We proposed that their attempt to compensate for their underperforming coparticipants during the practice trials may have trapped these participants into working hard on the final trial to appear consistent. Williams and Karau (1991) also found evidence for social compensation, but in collective rather than coactive settings. Although these findings may seem inconsistent, in both studies, participants worked hard only when they stood to gain something as a result. In the Williams and Karau study, participants could achieve a good collective performance with a low-effort partner only if they personally worked hard. In our Experiment 1, a good collective performance in the low-instrumentality condition was unlikely. Thus, in the no-evaluation condition, participants had no reason to work hard and, consequently, they loafed. In the evaluation condition, however, participants worked hard presumably to appear consistent to the experimenter and maintain a favorable experimenter evaluation.

An Alternative (Cost-Benefit) Model of Social Loafing

It is arguable that the expectancy-value model of low motivation and productivity in collective settings is needlessly complex and that social loafing can be explained more parsimoniously using a cost-benefit analysis (see Shepperd, 1993). That is, the expectancy and instrumentality components of the expectancy-value model are unnecessary for understanding low motivation and productivity in collective settings. Individuals will work hard provided that the benefits of high effort (e.g., achieving a valued outcome) outweigh the costs of high effort (e.g., exhaustion, depletion of personal resources, the annoyance of being a sucker to other's free-riding). According to this explanation, manipulations of expectancy and instrumentality affect contributing in collective settings by influencing the value that participants attach to their contributions. Moreover, manipulations of expectancy and instrumentality are actually manipulations of the value of the contributions. For example, when participants perceive that their personal contributions are dispensable in achieving a collective outcome, they likely regard their contributions as less valuable to the collective. A cost-benefit approach would argue that

it is ultimately the low value attached to contributing and not low expectations per se that produce a reduction in contributing when personal efforts are dispensable.

In many settings, expectations, instrumentality, and value are likely nonorthogonal, with changes in perceived expectations and instrumentality affecting the perceived value of the outcome or perceptions in perceived value affecting expectations and instrumentality. The present findings, nevertheless, argue against a strictly cost-benefits interpretation of social loafing. In both Experiments 1 and 2, the reward offered for a good performance was identical across the reward conditions. What varied in the reward conditions was the likelihood that a good collective performance would be achieved in Experiment 1 and the likelihood that a good collective performance would be rewarded in Experiment 2. In both experiments, participants worked hard only when instrumentality was high. Moreover, the effects of instrumentality were not affected by outcome value, that is, the manipulations of instrumentality continued to predict participant effort even after controlling for how much participants wanted to receive the reward. In short, our findings provide evidence that social loafing is better understood in terms of expectancy-value theory.

Conclusions

The present research manipulated aspects of instrumentality (the contingency between individual performance and group performance and the contingency between group performance and group outcome) while holding expectancy and value constant. We recognize, however, the importance of additional research that examines expectations, value (or costs associated with contributing), and other aspects of instrumentality. For example, a study of value would keep expectancy and instrumentality high while manipulating the value of the reward (either meager or large) offered for a good collective performance. We would predict that participants would work hard only when the reward offered for a good collective performance was large, exceeding the costs of contributing.

Low motivation and productivity seem endemic to settings in which individuals pool their efforts toward a common goal. Perhaps this should not be surprising given that individual efforts in collective settings are often unidentifiable and redundant and that the link between personal effort and achieving the desired collective outcome often is weak. The results from the present research add to the growing evidence that social loafing and free-riding can be reduced, perhaps even eliminated, if contributors to the collective (a) value the behavior or value the outcome to which they are contributing, (b) believe that performance is dependent on their efforts, and (c) believe that the outcome is dependent on performance.

NOTES

1. During pilot testing, we also varied the number of practice trials. When we had only one or two practice trials, participants typically displayed extremely high productivity in both the low-expectancy/evaluation and the low-expectancy/no-evaluation condition, consistent with research on social compensation (Williams & Karau, 1991). It was only after experiencing four practice trials that participants in the low-expectancy/no-evaluation condition (but not participants in the low-expectancy/evaluation condition) consistently reduced their effort, apparently reasoning that no amount of compensation would affect the group's chances of winning the prize.

2. During the development and piloting of the low-expectation manipulation, we found that participants were quite optimistic of their chances of winning the pizza (estimates averaging around 60%) even when objectively their chances were quite low (less than 2%). This led us eventually to present implicitly the probability of winning pizza as 1 in 200. Even so, low-expectancy participants rated their probability of winning the pizza as almost 1 in 4. As an aside, although we originally planned to run groups of five, last minute lab space limitations forced us to run groups of three. This resulted in an error in the percentages that we reported to participants. We reported that there would be at least 40 groups and 200 participants. However, because there were 3 participants per group, we should have reported a total 120 participants. Like ourselves, however, no participant seemed to notice this error during the experiment.

REFERENCES

Harkins, S. G., & Szymanski, K. (1988). Social loafing and self-evaluation with an objective standard. *Journal of Experimental Social Psychology, 24*, 354-365.

Ilgen, D. R., Nebeker, D. M., & Pritchard, R. D. (1981). Expectancy theory measures: An empirical comparison in an experimental simulation. *Organizational Behavior and Human Performance, 28*, 189-223.

Ingham, A. G., Levinger, G., Graves, J., & Peckham, V. (1974). The Ringelmann effect: Studies of group size and group performance. *Journal of Experimental Social Psychology, 10*, 371-384.

Karau, S. J., & Williams, K. D. (1993). Social loafing: A meta-analytic review and theoretical integration. *Journal of Personality and Social Psychology, 65*, 681-706.

Kerr, N. L. (1983). Motivation losses in small groups: A social dilemma analysis. *Journal of Personality and Social Psychology, 45*, 819-828.

Kerr, N. L. (1986). Motivational choices in task groups: A paradigm for social dilemma research. In H.A.M. Wilke, D. M. Messick, & C. G. Rutte (Eds.), *Experimental social dilemmas* (pp. 1-27). New York: Lang.

Kerr, N. L., & Bruun, S. E. (1983). Dispensability of member effort and group motivation losses: Free-rider effects. *Journal of Personality and Social Psychology, 44*, 78-94.

Latané, B., Williams, K., & Harkins, S. (1979). Many hands make light the work: The causes of consequences of social loafing. *Journal of Personality and Social Psychology, 37*, 822-832.

Mitchell, T. R. (1974). Expectancy models of job satisfaction, occupational preference and effort: A theoretical, methodological, and empirical appraisal. *Psychological Bulletin, 81*, 1053-1077.

Mitchell, T. R. (1982). Expectancy-value models in organizational psychology. In N. T. Feather (Ed.), *Expectations and actions: Expectancy-value models in psychology* (pp. 293-312). Hillsdale, NJ: Lawrence Erlbaum.

Mullen, B., & Baumeister, R. F. (1987). Group effects on self-attention and performance: Social loafing, social facilitation, and social impairment. In C. Hendricks (Ed.), *Review of personality and social psychology* (Vol. 9, pp. 189-206). Newbury Park, CA: Sage.

Peters, L. H. (1977). Cognitive models of motivation, expectancy theory and effort: An analysis and empirical test. *Organizational Behavior and Human Performance, 20*, 129-148.

Petty, R. E., Harkins, S. G., Williams, K., & Latané, B. (1977). The effects of group size on cognitive effort and evaluation. *Personality and Social Psychology Bulletin, 3,* 579-582.

Porter, L. W., & Lawler, E. E., III. (1968). *Managerial attitudes and performance*. Homewood, IL: Dorsey.

Pritchard, R. D., & De Leo, P. J. (1973). Experimental test of the valence-instrumentality relationship in job performance. *Journal of Applied Psychology, 57,* 264-270.

Sanna, L. J. (1992). Self-efficacy theory: Implications for social facilitation and social loafing. *Journal of Personality and Social Psychology, 62,* 774-786.

Sanna, L. J., & Pusecker, P. A. (1994). Self-efficacy, valence of self-evaluation, and performance. *Personality and Social Psychology Bulletin, 20,* 82-92.

Sanna, L. J., & Shotland, R. L. (1990). Valence of anticipated evaluation and social facilitation. *Journal of Experimental Social Psychology, 26,* 82-92.

Shepperd, J. A. (1993). Productivity loss in groups: A motivation analysis. *Psychological Bulletin, 113,* 67-81.

Shepperd, J. A. (1995). Remedying motivation and productivity loss in collective settings. *Current Directions in Psychological Science, 4,* 131-134.

Stroebe, W., & Frey, B. S. (1982). Self-interest and collective action: The economics and psychology of public goods. *British Journal of Social Psychology, 21,* 121-137.

Vroom, V. H. (1964). *Work and motivation*. New York: John Wiley.

Weldon, E., & Mustari, E. L. (1988). Felt dispensability in groups of coactors: The effects of shared responsibility and explicit anonymity on cognitive effort. *Organizational Behavior and Human Decision Processes, 41,* 330-351.

Williams, K. D., Harkins, S. G., & Latané, B. (1981). Identifiability as a deterrent to social loafing: Two cheering experiments. *Journal of Personality and Social Psychology, 40,* 303-311.

Williams, K. D., & Karau, S. J. (1991). Social loafing and social compensation: The effects of expectations of co-worker performance. *Journal of Personality and Social Psychology, 61,* 570-581.

Received October 7, 1997
Revision accepted June 14, 1998