

# Social Loafing and Self-Evaluation With a Social Standard

Kate Szymanski and Stephen G. Harkins  
Northeastern University

Social loafing has been described as the phenomenon in which participants who work together generate less effort than do participants who work alone (e.g., Latané, Williams, & Harkins, 1979). Subsequent research (Harkins & Jackson, 1985; Williams, Harkins, & Latané, 1981) has shown that a particular aspect of this paradigm leads to the loafing effect. When participants "work together," their outputs are pooled (combined) so that evaluation of individual output is not possible. In those studies, the evaluation potential of the experimenter has been emphasized. However, when the experimenter could not evaluate individual outputs, neither could the participants evaluate themselves. In this study we tested the possibility that the opportunity for the participants to evaluate themselves would be sufficient to eliminate the loafing effect. In two experiments, the evaluation potential of the experimenter (experimenter evaluation vs. no experimenter evaluation) was crossed with the potential for self-evaluation (self-evaluation vs. no self-evaluation). In both experiments, consistent with previous loafing research, the potential for evaluation by the experimenter was sufficient to increase motivation, whether participants could self-evaluate or not. However, when the experimenter could not evaluate the participants' outputs, the potential for self-evaluation reliably improved participant performance. In fact, self-evaluation was the only motivation needed for participants to exert as much effort as that exhibited by participants who could be evaluated by the experimenter.

Social loafing has been described as the phenomenon in which participants who work together generate less effort than do participants who work alone (Latané, Williams, & Harkins, 1979). This reduction in effort has been found to occur on tasks that require both physical effort (clapping, Harkins, Latané, & Williams, 1980; pumping air, Kerr & Bruun, 1981; shouting, Latané et al., 1979) and cognitive effort (reacting to proposals, Brickner, Harkins, & Ostrom, 1986; brainstorming and vigilance, Harkins & Petty, 1982; solving mazes, Jackson & Williams, 1985; evaluating essays, Petty, Harkins, Williams, & Latané, 1977). Social loafing has also been found to characterize the behavior of both females and males (Harkins et al., 1980), and in both between- and within-subjects designs (Harkins et al., 1980; Kerr & Bruun, 1981).

Although Latané et al.'s (1979) description implies that social loafing is simply a consequence of participants "working together," Harkins (1987) has suggested that this reduction in effort is a consequence of a particular feature of the experiments in which this effect has been found. In all of the loafing experiments, when participants worked together, their outputs were pooled (combined). Thus, individual outputs were "lost in the crowd," and participants could receive neither credit nor blame for their performances. Harkins (1987) has argued that it was the lack of evaluation potential that led to loafing in these experiments.

Consistent with this notion are the findings of Williams, Harkins, and Latané (1981; see also Kerr & Bruun, 1981), who led participants to believe that their outputs could always be monitored, whether they shouted individually or with others. Under

those conditions, participants worked as hard in groups as when alone. Also, when participants were led to believe that the experimenter's interest centered on group performance, and that individual outputs were to be pooled, the participants loafed as much when they performed individually as when they performed together. Williams et al. (1981) concluded, on the basis of these findings, that identifiability of individual outputs was an important mediator of social loafing.

However, Williams et al. (1981) did not manipulate only the identifiability factor. The participants all worked on exactly the same task; thus, when their performances were identifiable, they could also be directly compared with the performances of the other participants. This opportunity for comparison may have led participants to believe that their outputs could be evaluated, and it was this potential for evaluation, not only identifiability, that motivated performance. To test this notion, Harkins and Jackson (1985) manipulated identifiability and evaluation potential orthogonally, using a brainstorming task in which participants were asked to generate as many uses for an object as they could.<sup>1</sup> Consistent with previous loafing research (e.g., Harkins & Petty, 1982), participants generated more uses for an object when outputs were identifiable than when outputs were pooled. However, this difference emerged only when partici-

<sup>1</sup> It should be noted that the instructions given in this "brainstorming" experiment are not typical brainstorming instructions. Participants in loafing studies are not asked to be "free wheeling" (i.e., to produce wild and zany ideas; Maginn & Harris, 1980, p. 221). Instead, the participants were asked to generate as many uses as possible and not to be concerned about the quality of the uses, which could be ordinary or unusual. Given these instructions, participants in previous research (e.g., Harkins & Petty, 1982) have shown the typical loafing effect: Participants whose outputs were individually identifiable generated more uses than did participants whose outputs were pooled.

pants believed that their outputs could be evaluated by comparison with their co-worker's performance.

These findings suggest that identifiability, independent of other factors, does not eliminate social loafing. Participants must also expect that their outputs will be evaluated. Without this potential for evaluation, participants whose outputs were individually identifiable generated as little effort as those whose outputs were pooled. These data suggest that people are motivated by the potential for evaluation. Because evaluation is not possible when outputs are pooled, it is this aspect of working together that leads to social loafing.

Latané et al.'s (1979) description of social loafing may be considered generic. That is, any motivation loss in groups may be labeled social loafing. However, as previously noted, research conducted since the introduction of the term (e.g., Harkins & Jackson, 1985; Kerr & Bruun, 1981; Williams et al., 1981) suggests that social loafing stems from a particular aspect of the design of experiments in which this effect has been found: the fact that the pooling manipulation eliminates the possibility of evaluation. Kerr (1983) has conceptualized social loafing as being only one type of group motivation loss effect. In his research, he replicated the loafing effect and demonstrated two other types of group motivation loss effects: free-riding and sucker effects. Consistent with Kerr's (1983) analysis, the term *social loafing* will be used in its limited sense hereinafter.

Clearly, evaluation potential is central to this analysis of the loafing effect, but evaluation by whom? There are three potential sources of evaluation in the loafing paradigm: the experimenter, the coactor (or coactors), and the participant. In previous social loafing research, the role of the experimenter as evaluator has been emphasized. For example, Harkins et al. (1980) wrote

The results (social loafing) are easily explained by a minimizing strategy where participants are motivated to work only as hard as necessary to gain credit for a good performance or to avoid blame for a bad one. When the experimenter was unable to monitor individual outputs directly, performers sloughed off. (p. 464)

When manipulation checks have been used, only the participants' perceptions of the experimenter's ability to evaluate individual outputs have been assessed (e.g., Harkins & Jackson, 1985; Harkins & Petty, 1982; Jackson & Williams, 1985).

However, in this research, it appears that when the experimenter could not evaluate individual performances, neither could the coactors, nor the participant. For example, when participants shouted (Latané et al., 1979), no evaluation was possible because the masking noise eliminated the possibility that the participants could hear or be heard. After the session, if the participants had access to the scores, it was possible for them to compare (evaluate) individual performances, but they were as unable as the experimenter to decompose group shouts. Thus, when outputs were pooled, participants may have felt that they could not evaluate their own outputs, nor could these outputs be evaluated by their fellow participants.

Social loafing has been characterized as a "social disease" because it has "negative consequences for individuals, social institutions, and societies," resulting in "a reduction in human efficiency, which leads to lowered profits and lowered benefits for all" (Latané et al., 1979, p. 831). Given these potential consequences, it is important to determine exactly what factors un-

derlie the loafing effect. To this end, we focused on another potential source of evaluation, the *self*.

The motivational efficacy of the potential for self-evaluation is particularly interesting, given the types of tasks typically used in loafing research. That is, it is quite easy to imagine research participants eagerly testing their skills on tasks that appear interesting and challenging, whether or not someone else can monitor their performances. However, the tasks used in loafing research have been described as "noncompetitive, boring, tiring" (Harkins et al., 1980) and "unlikely to be personally involving for students, providing intrinsic importance, personal meaning, or significant consequences for one's life" (Brickner et al., 1986). In observations of the loafing effect, it may well have been the uninspiring nature of the tasks performed that led to emphasis on the importance of external evaluation (e.g., Harkins et al., 1980). However, in this research we tested the intriguing possibility that the opportunity for self-evaluation would be sufficient to eliminate the loafing effect, even when participants perform uninteresting tasks.

For evaluation by participants or by any other source to be possible, we would argue that two pieces of information must be known: the participant's individual output and a standard against which this output can be compared. The need for a standard can be satisfied in several ways. If a participant has had extensive experience with a given task, it is possible that he or she could have some notion of how his or her current level of performance compares with those earlier levels. In this case, the standard would be the participant's own earlier performances. However, in most, if not all, loafing research, it is unlikely that this type of standard would be available because the tasks are unfamiliar to the participants. In the absence of a personal standard, the need for a standard could also be satisfied by either an objective or a social criterion. If the task were optimizing (requiring some criterion performance; Steiner, 1972), both standards could be used. For example, if participants were asked to monitor a visual display and to report the number of signals they saw, the objective standard for this task would be the number of signals that were actually presented. For the same task, a social standard would be the average number of signals detected by previous participants. If the task were maximizing (requiring as much effort as possible), there would be no objective criterion for performance; only a social standard could be used. For example, if participants were asked to generate as many uses as possible for an object, the standard could be the average number of uses generated by previous participants.

To apply this analysis, it is also necessary to take into consideration what the participant believes he or she knows about his or her task performance. On some tasks, participants may believe that they know exactly what their output was, whereas on other tasks they may believe that they do not know what their output was. For example, pilot research has suggested that after taking part in the vigilance task, participants do not think that they know how many signals they detected. Conversely, preliminary research has shown that after participants have generated as many uses for an object as they can (brainstorming), they believe that they know how many uses they have generated. This suggests that the potential for self-evaluation on this task may be manipulated simply by providing or withholding a standard. Because this brainstorming task is maximizing, a social crite-

tion would be required (e.g., the average number of uses generated for this object by previous participants).

### Experiment 1

To determine whether providing a social standard was sufficient to give participants the impression that they could self-evaluate, we conducted a study in which a direct manipulation of the potential for experimenter evaluation was crossed with the manipulation of the potential for self-evaluation. Individual participants were asked to generate as many uses as they could for a knife. One half of the participants were informed that at the end of the session, the experimenter would count how many uses each person generated. The other one half of the participants were led to believe that the experimenter would not be able to evaluate individual performances. Crossed with this manipulation of experimenter evaluation was a manipulation of self-evaluation. One half of the participants were told that after they finished the task, they would be provided with the average number of uses generated in a previous experiment, whereas the other one half were told that to ensure confidentiality, this information would be withheld. We hypothesized that, in accordance with previous research on social loafing (Williams et al., 1981), when there was no opportunity for self-evaluation, participants who anticipated experimenter evaluation should generate more uses than would participants who believed that the experimenter could not evaluate their individual efforts. However, if the manipulation of the potential for self-evaluation was successful, and this possibility motivated performance, the no experimenter evaluation/self-evaluation participants should generate more uses than would the no experimenter evaluation/no self-evaluation participants.

### Method

**Participants.** Sixty women and men took part in this study as a means of fulfilling an introductory psychology course requirement. In this study, experimenter evaluation vs. no experimenter evaluation was crossed with self-evaluation vs. no self-evaluation. Participants were tested individually.

**Procedure.** Participants were seated at a desk, next to which a cardboard box was placed. They were informed that they would be given the name of an object and their task would be to generate as many uses as possible for this object. They were not to be concerned about whether the uses were ordinary or unusual, but to generate as many uses as possible. Participants were provided with an envelope that contained pieces of paper on which they were to write their uses. They were asked to write one use per slip, to fold the slip, to set it aside, and then to write the next use on a second slip, and so on. Participants were informed that a bell would sound when they were to stop.

Next, the participants were informed that to protect their interests, we were attempting to ensure the confidentiality of their performances in any way we could. In the experimenter evaluation condition, participants were told

To ensure confidentiality, on our research team only the experimenter will know how many uses you generated. After the bell sounds, the experimenter will count the number of uses you generated, and then add them to those generated by the other participants in the box you see.

In the no experimenter evaluation condition, the participants were told

To ensure confidentiality, no one on our research team will know how many uses you generated. After the bell sounds, add your uses to those generated by the other participants in the box that you see.

All of the participants were shown a box that contained a number of slips, ostensibly generated by the preceding participants. All participants were then informed that we were interested in determining the average performance on this task, and that at the end of the experiment, we would calculate this average by adding the total number of uses generated and dividing by the total number of participants. We were interested in comparing this average with the average number generated by students in previous research.

Crossed with the manipulation of experimenter evaluation was the manipulation of self-evaluation. Participants in the self-evaluation condition were told

Since we thought you might be interested, at the end of your session we will tell you the average number of uses generated by students in this other research. However, please hold this information in confidence since it could affect the performance of later participants.

In the no self-evaluation condition, the participants were told

To ensure confidentiality, we will not be able to give you information about the average number of uses generated in this other replication since this information could affect the performance of later participants.

After having heard their respective instructions, participants were given twelve minutes to generate uses for a knife. After that time, they were asked to fill out a questionnaire that included manipulation checks. When they completed the questionnaire, participants were dismissed.

Unbeknownst to the participants, there were exactly 70 slips of paper in the envelopes that they were given at the beginning of the experiment. By counting the number of slips left by each person, we determined the number of uses generated by each participant, even in the no experimenter evaluation condition.

### Results

The data were analyzed in 2 (experimenter evaluation vs. no experimenter evaluation)  $\times$  2 (self evaluation vs. no self evaluation) analyses of variance (ANOVAS). Where appropriate, the Newman-Keuls (Kirk, 1982) was used as a post hoc test to compare means.

**Manipulation checks.** Participants were asked to what extent the experimenter would be able to evaluate their individual performances. Analysis of the participants' answers revealed that participants in the experimenter evaluation condition believed to a greater extent that the experimenter would be able to evaluate their performances ( $M = 8.9$  on an 11-point scale) than did participants in the no experimenter evaluation condition ( $M = 2.4$ ),  $F(1, 64) = 138.9$ ,  $p < .001$ . There were no other reliable effects on this measure ( $ps > .20$ ). Participants were also asked to what extent they would be able to know exactly how well they performed individually. Analysis of their answers revealed that participants in the self-evaluation condition believed that they could evaluate their own performances to a greater extent ( $M = 7.8$ ) than could no self-evaluation participants ( $M = 3.4$ ),  $F(1, 64) = 45.1$ ,  $p < .001$ . No other effects approached significance ( $ps > .20$ ).

**Uses.** Analysis of the uses data revealed an Experimenter Evaluation  $\times$  Self-Evaluation interaction,  $F(1, 64) = 11.1$ ,  $p <$

.001. Experimenter evaluation led to equivalent performances in the self-evaluation ( $M = 23.6$ ) and no self-evaluation conditions ( $M = 27.0$ ,  $p > .20$ ). The participants in each of these conditions generated more uses than did participants in the no experimenter evaluation/no self-evaluation condition ( $M = 16.8$ ,  $ps < .05$ ). When the experimenter could not evaluate the participants' performances, but self-evaluation was possible, participants also generated more uses ( $M = 28.6$ ) than were produced when no self-evaluation was possible ( $p < .01$ ). In fact, self-evaluation alone motivated participants to generate as many uses as were generated in the experimenter evaluation conditions ( $ps > .20$ ).

### Discussion

We argued that previous research has shown that the potential for evaluation is central to the loafing effect (e.g., Williams et al., 1981; Harkins & Jackson, 1985), but previous research has not specified which of the potential sources is responsible for the effect. Experiment 1 tested one method of manipulating the potential for self-evaluation. We hypothesized that two pieces of information are necessary for evaluation by any of the sources to be possible: a participant's individual output and a standard. Preliminary research on the brainstorming task indicated that participants believed that they knew how many uses they had generated. This suggests that to manipulate the potential for self-evaluation, the experimenter need only provide or withhold a standard. Because the brainstorming task used in this study was maximizing, we used a social standard: the average number of uses generated in a previous experiment. Analysis of the participants' perceptions of their ability to self-evaluate revealed that this manipulation was successful. Participants who were told that they would be given the average from a previous experiment indicated, to a greater extent than did participants who were not given this cue, that they would know exactly how well they performed. The manipulation of experimenter evaluation was also successful. Participants who added their slips to the box before the experimenter entered the room reported that the experimenter was less able to tell exactly how well they performed than did participants whose uses were counted by the experimenter before they were added to the box.

Analysis of the number of uses generated for the object revealed that the potential for experimenter evaluation motivated performance, regardless of the potential for self-evaluation. Participants in the no experimenter evaluation/no self-evaluation condition generated fewer uses than did participants in either of the experimenter evaluation conditions. In previous loafing research, it has been argued that participants loaf because an external source, the experimenter, cannot evaluate their performances (Harkins et al., 1980). This condition was met in the no experimenter evaluation/self-evaluation condition; nonetheless, these participants, given only the opportunity for self-evaluation, generated as many uses for the object as did participants who could be evaluated by the experimenter.

These results suggest that this method of manipulating the potential for self-evaluation was successful and that this potential also motivated performance. However, this experimental paradigm is different from the typical loafing paradigm. First, participants were tested individually instead of in groups. Second, in the no experimenter evaluation conditions, the partici-

pants' outputs were added to the uses ostensibly generated by previous participants, rather than to the uses generated by coparticipants, as has usually been the case in loafing research (e.g., Harkins & Jackson, 1985). Although the potential for self-evaluation motivated performance in Experiment 1, it is possible that working in the physical presence of others creates a different psychological environment for the participants. For example, when participants are physically present, the notion that responsibility for the total output is shared may be more salient than when the participant works individually. Under these circumstances, the opportunity for self-evaluation may not be sufficient to motivate performance. In addition, although data have already been collected that suggest that participants in the pooled output condition feel that there is little opportunity for evaluation by the experimenter (e.g., Harkins & Jackson, 1985; Harkins & Petty, 1982; Jackson & Williams, 1985), there is no direct evidence that participants in these pooled output conditions also feel that there is reduced opportunity for self- or coparticipant evaluation. Experiment 1 is not informative in this regard, because participants were tested individually.

### Experiment 2

To address these issues, as well as to demonstrate the replicability of the self-evaluation effect, we conducted a second experiment in which we used the typical loafing paradigm. Participants were tested in pairs and were asked to generate as many uses as possible for a new object, a box. For one half of the pairs, the participants' uses were collected in separate bins (individually identifiable) in a collection box, whereas the uses for the other one half were collected (pooled) in a single bin. Crossed with this manipulation, one third of the pairs were told that this experiment was a replication of a previous experiment and that at the end of the session they would be told how many uses, on average (standard), had been generated in that earlier experiment; one third of the pairs were told that this was a replication of a previous experiment, but that they could not be told the average number of uses because that information could affect the performance of later participants if it were known (no standard); and finally, one third of the pairs were told nothing about a previous experiment, and served as a social loafing control condition.

To sustain our argument concerning evaluation potential in previous loafing research, we hypothesized that participants in the pooled/control condition should indicate that none of the potential sources (experimenter, coparticipants, self) could evaluate individual performances. If the experimenters provide the average *number of uses*, the pooled/standard participants should be prompted to indicate that they can self-evaluate. In replication of Experiment 1's findings, this possibility should be sufficient to motivate performance, despite the fact that the opportunity for evaluation by the experimenter and coactor was minimized.

### Method

**Participants.** As a means of fulfilling an introductory psychology course requirement, 144 women and men took part in this experiment. Participants were tested in pairs; each pair was randomly assigned to one of six conditions comprising a 2 (individually identifiable vs. pooled)  $\times$  3 (standard, no standard, control) design.

*Procedure.* Participants were seated at a desk and separated from each other by a partition. A covered box was placed on the floor in front of the desk. Two tubes extended from the box to the desk, one for each participant. Participants were informed that they would be given the name of an object and that their task would be to generate as many uses for this object as they could. These uses could be ordinary or unusual; they were simply to come up with as many uses as possible. Each use was to be written on a separate slip of paper that was then folded and slipped down the tube. They were informed that a bell would sound when they were to stop.

Next, participants in the individually identifiable condition were told that the experimenter was interested in the total number of uses generated by each individual and were shown that the collection box was divided, so that at the end of the session each person's output could be counted by the experimenter.

In the pooled condition, the participants were informed that the experimenter was interested in the total number of uses generated by the group and were shown that there was no divider in the box, so that their uses would be pooled. At the end of the session, the experimenter would count the total number of uses generated by each pair and determine how well that particular pair performed.

Crossed with this manipulation of evaluation potential, one third of the participants were told that this was a replication of an experiment conducted in the previous school quarter and that we were interested in seeing whether or not we could replicate those findings. We went on to say

Since we thought you might be interested, at the end of your session, we will tell you the average number of uses generated by students in this other research. However, please hold this information in confidence since it could affect the performance of later participants (standard).

Another one third of the pairs were also told that the current research was a replication of a previous experiment.

However, to ensure confidentiality, we will not be able to give you the average number of uses generated by students in this previous research since this information could affect the performance of later participants (no standard).

The final one third of the pairs were told nothing about the previous experiment, so no information about a standard was mentioned (social loafing control).

After having heard their respective instructions, participants had 12 min to generate uses for a box. After that time, they were asked to fill out a questionnaire that included manipulation checks. When they completed the questionnaire, the participants were dismissed.

## Results

The responses for each pair were averaged, so the group served as the unit of analysis. These data were analyzed in 2 (identifiable vs. pooled)  $\times$  3 (standard, no standard, control) ANOVAs, and where appropriate, the Newman-Keuls (Kirk, 1982) was used as an a posteriori test.

*Manipulation checks.* Participants whose outputs were identifiable acknowledged that the experimenter would be better able to evaluate their performances ( $M = 8.4$  on an 11-point scale) than would participants whose outputs were pooled ( $M = 4.8$ ),  $F(1, 66) = 43.4$ ,  $p < .001$ . Analysis of a question asking participants to rate to what extent they would be able to evaluate their own performances yielded an instruction main effect,  $F(2, 66) = 21.0$ ,  $p < .001$ . Participants who were provided a standard reported that they would be better able to evaluate their own performances ( $M = 7.7$ ) than would participants

who were given no standard instructions ( $M = 4.5$ ) or the control instructions ( $M = 4.8$ ,  $ps < .05$ ). The means for the latter two conditions did not differ ( $p > .20$ ). It should be noted that in the absence of any information, participants in the control condition did not feel that they could self-evaluate any more than could no standard participants who were explicitly told that they would not be given a standard. No other effects were reliable on either of these measures ( $ps > .20$ ). Finally, participants in all conditions felt that they could not be evaluated by their coparticipants (overall  $M = 3.2$ ).

*Uses.* Analysis of the number of uses generated by the participants revealed two main effects and an interaction. Participants whose outputs could be evaluated by the experimenter generated more uses ( $M = 26.6$ ) than did those whose outputs were pooled ( $M = 21.8$ ),  $F(1, 66) = 14.7$ ,  $p < .001$ . Participants given standard instructions generated reliably more uses ( $M = 26.6$ ) than did participants in the no standard conditions ( $M = 21.8$ ,  $p < .05$ ). The number of uses generated by participants in the control conditions ( $M = 24.7$ ) did not differ reliably from the number generated in the other two conditions ( $ps > .20$ ),  $F(2, 66) = 5.0$ ,  $p < .05$ .

More interesting is the two-way interaction,  $F(2, 66) = 6.2$ ,  $p < .004$ . In the replication of previous social loafing research that used this task, participants in the control conditions whose outputs could be evaluated by the experimenter generated more uses ( $M = 28.5$ ) than did participants whose outputs were pooled ( $M = 20.0$ ,  $p < .05$ ). Similarly, when given no standard instructions, participants whose outputs could be evaluated generated more uses ( $M = 25.4$ ) than did participants whose outputs were pooled ( $M = 18.2$ ,  $p < .05$ ). However, when participants were provided a standard, participants whose outputs could not be evaluated by the experimenter generated as many uses ( $M = 27.3$ ) as did participants subject to external evaluation ( $M = 25.9$ ,  $p > .20$ ).

## Discussion

Previous loafing research has emphasized the role of the experimenter as evaluator (e.g., Harkins et al., 1980). However, we have argued that when outputs are pooled, participants may feel that not only is the experimenter unable to evaluate their individual outputs, but also, evaluation is not possible by the participants and coactors. In Experiment 2, we tested this notion by asking the participants to rate the extent to which they believed that evaluation by each of the potential sources (experimenter, coactor, self) was possible. Consistent with this analysis, participants in the pooled/control condition reported that it was unlikely that they could be evaluated by the experimenter ( $M = 5.1$ ), by their coactor ( $M = 3.2$ ), or that they could evaluate themselves ( $M = 4.4$ ). These results suggest that participants may loaf, not only because they feel that evaluation by the experimenter is not possible, but also because they feel that they cannot self-evaluate, nor can they be evaluated by their coactors.

In Experiment 2, we tested the possibility that the opportunity for self-evaluation would be sufficient to motivate performance in the typical loafing paradigm. As in Experiment 1, the opportunity for self-evaluation was manipulated by providing or withholding the average number of uses generated in a previous version of the experiment. Analysis of the manipulation

checks revealed that providing this standard was sufficient to lead participants in the standard condition to report that they could self-evaluate to a reliably greater extent than could participants in the no standard or control conditions.

In these experiments, we simply provided participants with the opportunity to self-evaluate, without attempting to motivate them to do so. Previous descriptions of the tasks used in loafing research (Brickner et al., 1986; Harkins et al., 1980) would suggest that taking part in these tasks would be sheer drudgery and that in the absence of external evaluation pressures, there would be little reason to perform well since the tasks are devoid of inherent interest. Nonetheless, the opportunity for self-evaluation was sufficient to lead to performance equivalent to that achieved in the experimenter evaluation conditions. However, even in this motivationally barren setting, the opportunity for self-evaluation actually provided the participants with two incentives: They could learn something about their abilities on this task and they also could take pleasure in surpassing the performance of the "average" previous participants. Thus, the participants could have been motivated by the possibilities for self-validation (learning that they were better than average), and increasing self-knowledge (learning something about their abilities on the task), or by either of the two possibilities. These two incentives will always be available when a social standard is used. With *maximizing* tasks in which participants have had no extensive prior experience, a social standard is all that is available. After all, there is no "correct" or "perfect" performance when one is asked to maximize one's efforts. However, on optimizing tasks, which require some criterion performance, both social and objective standards are available. For example, on a vigilance task, which requires the participants to report seldomly occurring signals, either type of standard could be used. That is, participants could be told how many signals previous participants had detected (social standard), how many signals had actually been presented (objective standard), or both. By using an optimizing task and only an objective standard, the possibility of self-validation could be removed, allowing us to determine whether the self-knowledge aspect is sufficient to eliminate the loafing effect.

The argument implicit in this analysis is that the opportunity for self-evaluation may be sufficient to increase motivation. We have bolstered this claim by pointing out how uninteresting are the tasks used in loafing research; nonetheless, people work on them, motivated only by the potential for self-evaluation. Using an optimizing task with an objective standard would carry this argument one step further by showing that the effect persists even when the possibility of self-validation is eliminated. However, simply having the opportunity to self-evaluate may not guarantee the motivation needed to do so. Obviously, some level of motivation must exist, and these experiments only begin to specify what is necessary to arouse and to maintain it. For example, in our experiments, the participants were quite likely to be performing the task for the first time. After they determined their level of proficiency, participants may have had little interest in spending additional energy on the task. If exposed to repeated trials, participants may have taken advantage of the opportunity to loaf. This outcome would be consistent with the argument that participants were motivated by the potential for self-evaluation, but once they had gained some notion of their competency on the task, participants would have felt no need

to maintain their performance. This suggests that the self-evaluation manipulation may have limited value as a mode of eliminating social loafing. On the other hand, small variations in the descriptions of the tasks (e.g., "This object may be more difficult to generate uses for than was the previous one") may be sufficient to maintain performance levels, because this may lead participants to believe that there is something more to be learned about their proficiency at the task.

Another factor that may affect the motivation to self-evaluate is the intrinsic appeal of the task. That is, if the tasks were more interesting, there may be greater motivation to determine one's competency at them. However, there are data which suggest that concern about self-evaluation may undermine motivation on tasks that have intrinsic appeal. Bartis, Szymanski, and Harkins (in press) have found that participants who were asked to come up with uses that were as creative as possible generated uses that were more creative when the uses could not be evaluated by anyone than when they could be evaluated by the experimenter. These findings are consistent with other research (e.g., Amabile, 1979) which suggests that on interesting, challenging tasks, the possibility of external evaluation undermines intrinsic motivation. It remains to be seen whether the potential for self-evaluation would have a similar effect, but these manipulations are similar in that they orient the participants toward achieving some goal (i.e., standard), rather than performing the task for its own sake.

Of course, these possibilities represent only a few of the ways to proceed in an attempt to determine the limits of the self-evaluation effect. Pursuing these possibilities may be worthwhile, not only because we will learn something more about social loafing, but also because we will learn something more about the motivational properties of the potential for self-evaluation, a notion that many social psychological theories have incorporated. For example, Festinger (1954) proposed that there exists a "drive" for people to evaluate their opinions and abilities. When objective means exist by which these judgments may be made, they will be used; however, this is often not the case. Under these circumstances, people will turn to comparisons with the opinions or abilities of others. In recent years, attention has been directed toward self-evaluation as part of the renewed interest in the self (e.g., Schlenker, 1985). For example, Greenwald and his colleagues (e.g., Greenwald & Breckler, 1985) proposed an ego-task analysis that specifies three facets of the self that are relevant for evaluation concerns: the public self, the private self, and the collective self. At the heart of this typology is the notion that behavior is motivated, in part, by concern for evaluation by these sources. Deci and Ryan (1985) developed a theory of intrinsic motivation and self-determination, a central tenet of which is that man possesses an innate organismic need for competence that energizes a variety of behaviors. As part of his grand theory of motivation, Bandura (1986) included an account of the process of self-evaluation that includes self-observation, judgmental processes, and self-reactions. The self-evaluation process affects judgments of self-efficacy that, in turn, affect a whole range of behaviors.

The current line of research provides data that are relevant for each of these accounts of self-evaluation effects. For example, most research that tests Festinger's (1954) theory has focused on the characteristics of the others chosen for comparison (e.g., Levine & Moreland, 1986). Our research provides infor-

mation about an earlier point in the process. That is, the theory proposes that there exists in the human organism a "drive" to evaluate one's opinions and abilities. Although the notion of a "drive" may be dated, the notion that people learn that it is useful to have a fairly accurate idea of their talents and skills remains quite current (e.g., Goethals & Darley, 1987). Our research provides data that are entirely consistent with this assumption, which lies at the heart of the theory. As another example, Greenwald and Breckler (1985) argued: "The prevalent assumption heretofore has been that self-presentations are targeted at audiences of others. We have reviewed evidence that there is also an important inner audience, oneself" (p. 141). Our findings are consistent with Greenwald and Breckler's (1985) notions concerning the importance of the private self. After all, in the crucial conditions in this research, the only person who will know how a given participant performed is that person who is doing the performing.

This research focused on the role that evaluation plays in producing loafing effects. We do not propose that effects stemming from manipulations of evaluation potential account for all, or even most motivation losses in groups. Any number of other variables may affect performance in group settings (e.g., dispensability of member effort; Kerr & Bruun, 1983). Rather, we are arguing that evaluation potential plays a central role in producing the reduction in effort that has been termed *social loafing*. Even when our attention is limited to the loafing paradigm, it is clear that other factors motivate performance, regardless of the potential for evaluation. For example, creativity (Bartis, Szymanski, & Harkins, in press), personal involvement (Brickner et al., 1986), task challenge and task uniqueness (Harkins & Petty, 1982), partner effort (Jackson & Harkins, 1985), and group cohesion (Williams, 1981) have all been shown to eliminate the loafing effect even though the potential for evaluation by each of the potential sources was minimized. In future research it will be necessary to determine how these other factors interact with the potential for evaluation to motivate performance in these settings.

## References

- Amabile, T. (1979). Effects of external evaluation on artistic creativity. *Journal of Personality and Social Psychology*, 37, 221-233.
- Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory*. Englewood Cliffs, NJ: Prentice-Hall.
- Bartis, S., Szymanski, K., & Harkins, S. (in press). Evaluation of performance: A two-edged knife. *Personality and Social Psychology Bulletin*.
- Brickner, M., Harkins, S., & Ostrom, T. (1986). Personal involvement: Thought provoking implications for social loafing. *Journal of Personality and Social Psychology*, 51, 763-769.
- Deci, E., & Ryan, R. (1985). *Intrinsic motivation and self-determination in human behavior*. New York: Plenum.
- Festinger, L. (1954). A theory of social comparison processes. *Human Relations*, 7, 117-140.
- Goethals, G., & Darley, J. (1987). Social comparison theory: Self-evaluation and group life. In B. Mullin & G. Goethals (Eds.), *Theories of group behavior*. (pp. 21-47). New York: Springer-Verlag.
- Greenwald, A., & Breckler, S. (1985). To whom is the self presented? In B. Schlenker (Ed.), *The self and social life*. (pp. 126-145). New York: McGraw-Hill.
- Harkins, S. (1987). Social loafing and social facilitation. *Journal of Experimental Social Psychology*, 23, 1-18.
- Harkins, S., & Jackson, J. (1985). The role of evaluation in eliminating social loafing. *Personality and Social Psychology Bulletin*, 11, 457-465.
- Harkins, S., Latané, B., & Williams, K. (1980). Social loafing: Allocating effort or taking it easy? *Journal of Experimental Social Psychology*, 16, 457-465.
- Harkins, S., & Petty, R. (1982). Effects of task difficulty and task uniqueness on social loafing. *Journal of Personality and Social Psychology*, 43, 1214-1229.
- Jackson, J., & Harkins, S. (1985). Equity in effort: An explanation of the social loafing effect. *Journal of Personality and Social Psychology*, 49, 1119-1206.
- Jackson, J., & Williams, K. (1985). Social loafing on difficult tasks: Working collectively can improve performance. *Journal of Personality and Social Psychology*, 49, 937-942.
- Kerr, N. (1983). Motivation losses in small groups: A social dilemma analysis. *Journal of Personality and Social Psychology*, 45, 819-828.
- Kerr, N., & Bruun, S. (1981). Ringelmann revisited: Alternative explanations for the social loafing effect. *Personality and Social Psychology Bulletin*, 7, 224-231.
- Kerr, N., & Bruun, S. (1983). Dispensability of member effort and group motivation losses: Free-rider effects. *Journal of Personality and Social Psychology*, 44, 78-94.
- Kirk, R. (1982). *Experimental design*. Belmont, CA: Wadsworth.
- Latané, B., Williams, K., & Harkins, S. (1979). Many hands make light the work: The causes and consequences of social loafing. *Journal of Personality and Social Psychology*, 37, 823-832.
- Levine, J., & Moreland, R. (1986). Outcome comparisons in group contexts: Consequences for the self and others. In R. Schwarzer (Ed.), *Self-related cognitions in anxiety and motivation*. (pp. 285-303). Hillsdale, NJ: Erlbaum.
- Maginn, B., & Harris, R. (1980). Effects of anticipated evaluation on individual brainstorming performance. *Journal of Applied Psychology*, 65, 219-225.
- Petty, R., Harkins, S., Williams, K., & Latané, B. (1977). The effects of group size on cognitive effort and evaluation. *Personality and Social Psychology Bulletin*, 3, 579-582.
- Schlenker, B. (1985). *The self and social life*. New York: McGraw-Hill.
- Steiner, I. (1972). *Group processes and productivity*. New York: Academic Press.
- Williams, K. (1981, May). *The effects of group cohesiveness on social loafing*. Paper presented at the meeting of the Midwestern Psychological Association, Detroit, MI.
- Williams, K., Harkins, S., & Latané, B. (1981). Identifiability as a deterrent to social loafing: Two cheering experiments. *Journal of Personality and Social Psychology*, 40, 303-311.

Received December 24, 1986

Revision received May 10, 1987

Accepted May 16, 1987 ■