Social Loafing and Social Facilitation

STEPHEN G. HARKINS

Northeastern University

Received May 29, 1986

Social facilitation and social loafing have been treated as separate lines of research in the social psychological literature. However, it is argued in the present paper that these two paradigms are closely related; in fact, they are complementary. Viewed from this perspective, the experimental conditions that have been included in loafing and facilitation research fall into three cells of a 2 (Alone vs. Coaction) × 2 (Evaluation vs. No Evaluation) factorial design. In the current research, the complete 2 × 2 design was run in two experiments. In both experiments, consistent with the findings of previous loafing research, with number held constant, participants whose outputs could be evaluated outperformed participants whose outputs could not be, but, inconsistent with descriptions of the loafing effect (e.g., B. Latané, K. Williams, & S. Harkins, 1979, Journal of Personality and Social Psychology, 37, 823–832), with evaluation potential held constant, pairs outperformed singles. These data suggest that both social facilitation and social loafing can be accommodated in the same design. It is argued that combining the paradigms in this way refines our understanding of both phenomena. © 1987 Academic Press, Inc.

In the first published experiment in social psychology, Triplett (1898) found that children reeled more fishing line when working alongside another child similarly occupied than when reeling alone. Some 15 years later, Ringelmann (1913, summarized by Kravitz & Martin, 1986) reported that students working together pulled on a rope with less force than was expected on the basis of their individual outputs. Reeling line and pulling rope are both very simple tasks. Yet on one, working together led to better performance than working alone, while on the other, the opposite effect was obtained. These two experiments represent the initial efforts in what have come to be considered two separate lines of research, social facilitation and social loafing. The fact that these two phenomena fall into separate research domains may account for the fact that nothing has been made of the apparently contradictory nature of the findings. However, we will argue that these findings are intimately related, and

Requests for reprints should be sent to Dr. Stephen G. Harkins, Department of Psychology, Northeastern University, 360 Huntington Avenue, Boston, MA 02115.
are not contradictory. In fact, the two research paradigms are complementary.

Actually, these experiments were seen at one time as falling in the same domain. In an early review, Dashiell (1935) included Ringelmann's study in a section entitled "The effects of coworkers upon the individual's work" along with other experiments on social facilitation (e.g., Allport, 1920). Given the inconsistencies that already characterized the research in this area, the discrepancy between the findings of Triplett and Ringelmann must not have appeared noteworthy. As Dashiell (1935) wrote: "Looking backward over this section it appears that results with tests on the effect of the presence of other workers alongside are not in agreement on all points. Very generally there seems to emerge a recognition of contrasted kinds of influence from coworkers, facilitating and inhibiting" (p. 1115). Matters had not progressed by the 1950s when Solomon Asch (1952) noted that so-called explanations for the effects of the presence of others on performance represented restatements of the results, explaining nothing. After several decades of research, there was no convincing refutation of Asch's suspicion, and interest in the area waned.

In 1965, Zajonc suggested a resolution to the muddle. He hypothesized that the "mere presence" of others leads to increased drive, which enhances the tendency to emit dominant responses. If the dominant response is correct, facilitation is obtained. If incorrect, performance is debilitated. Of course, on simple or well-learned tasks, the dominant response is likely to be the correct one, leading to facilitated performance, as in Triplett's case. However, on complex or not well-learned tasks, the dominant response is likely to be incorrect, leading to debilitated performance. Zajonc's (1965) review suggested that this analysis made sense of a large body of research.

But what of Ringelmann's finding? The task was quite simple and people worked together, yet, there was no facilitation effect. Ringelmann's research was not included in Zajonc's review, apparently because, by now, it was considered to be part of the group process literature. In Zajonc's (1966) text, social facilitation effects were presented in a chapter entitled "Coaction," while Ringelmann's findings appeared in the "Group Performance" chapter, and why not? In Triplett's study, "working together" meant working individually, side by side, on the same task. Ringelmann's students actually "worked together," pulling on the same rope. In the latter case, as Zajonc (1966) noted, the participants had to coordinate their efforts to achieve their full potential. If the participants reached their peak pulls at slightly different times, or if they pulled along slightly different axes, coordination loss would occur, leading to suboptimal performance. As Steiner (1972) has shown, the performance decrements exhibited by Ringelmann's students were directly proportional to the number of coordination links among the participants. The potential for
coordination loss is clearly a consequence of working with others. Therefore, it makes sense to consider the Ringelmann effect as a "Group Process," while Triplett's study and the subsequent research are placed in the domain of "Coaction" (i.e., the effects of the presence of others working independently on the same task). Following this classification scheme, Forsyth (1983) presented coaction research in a section of his book entitled "Performance when others are present," while social loafing was presented in "Performance in interacting groups." In most social psychology textbooks, these two phenomena appear in separate sections and no consideration is given to their relationship.

Subsequent research, however, has shown that coordination loss alone is insufficient to account for Ringelmann's findings, and that working together can lead to less output even when there is no opportunity for coordination loss. Ingham, Levinger, Graves, and Peckham (1974), after replicating Ringelmann's rope-pulling experiment, arranged things in a second experiment so that, on certain trials, participants were given the impression that they were pulling with others, when they actually pulled alone. This arrangement provided no opportunity for coordination loss. On these pseudogroup trials, participants put out less effort than when they thought they were pulling alone. This reduction in effort has been termed social loafing (Latané, Williams, & Harkins, 1979), and this effect has been demonstrated for both sexes on tasks requiring 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; brainstorming1 and vigilance, Harkins & Petty, 1982; solving mazes, Jackson & Williams, 1985; evaluating essays, Petty, Harkins, Williams, & Latané, 1977).

Applying Steiner's (1972) typology, some of these tasks were "maximizing" (requiring the participant to put out as much effort as possible, such as rope-pulling, shouting, pumping air, and brainstorming), while others were "optimizing" (requiring the participant to achieve some criterion performance, such as evaluating essays, vigilance, and solving mazes). On all of these tasks, "groups" failed to achieve even the potential suggested by their individual performances, though the possibility of coordination loss was eliminated, either by using pseudogroups (e.g., Ingham et al., 1974; Latané et al., 1979), or by using tasks on which

---

1 It should be noted that the instructions given in the "brainstorming" experiment are not typical brainstorming instructions. Participants in loafing studies were 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 participants whose outputs were pooled.
individual performances added, with no possibility of coordination loss (e.g., vigilance and brainstorming, Harkins & Petty, 1982; solving mazes, Jackson & Williams, 1985).

Thus, while the potential for coordination loss may have consigned Ringelmann’s findings to the group domain, this difference between the paradigms is not sufficient to account for the contradictory findings. In fact, a close examination of the two experimental paradigms reveals many more similarities than differences.

In the prototypic coaction experiment, the outputs of the coacting participants are individually identifiable, and these outputs are compared to the output of a single participant performing the same task. The usual finding in social facilitation research is that working together leads to enhanced performance on simple tasks and debilitated performance on complex ones. As noted previously, Zajonc (1965) offered a drive interpretation of this phenomenon. Over the next decade, this drive account was seen as the most parsimonious theoretical explanation for facilitation effects (e.g., Geen & Gange, 1977), although there was controversy over whether it was the mere presence of the others that increased drive (Zajonc, 1965, 1980), or the fact that these others were associated with evaluation and/or competition (Cottrell, 1972).

However, in a review of the facilitation literature since 1977, Geen (in press) concluded that “today such a confident assertion of the primacy of the drive theoretical approach is not warranted.” “Instead several sophisticated alternatives have found considerable support in experimental studies.” Geen (in press) organized these theoretical approaches into three broad classes. One class of theories incorporated those approaches that continue to rely on the notion that the presence of others increases drive (e.g., distraction/conflict, Baron, 1986; evaluation apprehension, Cottrell, 1972; social monitoring, Guerin & Innes, 1982; compresence, Zajonc, 1980). The second class of theories included those approaches that suggest that “the presence of others creates either explicit or implicit demands on the person to behave in some way” (Geen, in press) (e.g., self-presentation, Bond, 1982; self-awareness, Carver & Scheier, 1981), while the third class consisted of Baron’s (1986) information processing view of distraction/conflict, which proposes that the presence of others affects focus of attention and information processing.

Although there are a number of theories that attempt to account for facilitation effects, and “none appears to command the high ground as drive theory did in the 1960s and 1970s” (Geen, in press), each of these theories focuses on the effects of one or both of the same two features of the facilitation paradigm: the mere presence of others, and/or the potential for evaluation that these others represent. The “number” theories argue that: the mere presence of others leads to the feeling of uncertainty which leads to increased drive (Zajonc, 1980); the presence of others
leads to uncertainty and increased drive, when the behavior of these others cannot be monitored (Guerin & Innes, 1982); the simple presence of others may be distracting enough to trigger attentional conflict and increased drive (Baron, 1986); or the simple presence of others may be distracting enough to create an attentional overload affecting focus of attention (Baron, 1986). The “evaluation” theories argue that: the presence of others can come to be associated with evaluation and/or competition, and the resulting evaluation apprehension leads to increased drive (Cottrell, 1972); the possibility of evaluation leads to distraction which increases drive (Baron, 1986); the potential for evaluation makes one self-aware which leads to greater attention to how performance matches some standard (Carver & Scheier, 1981); the prospect of evaluation leads to concerns about self-presentation (Bond, 1982); or the potential for evaluation affects focus of attention (Baron, 1986). Thus, each of these explanations represents an attempt to account for the effects of number and/or evaluation.

Social loafing has been described as the finding that participants working together put out less effort than participants working alone (e.g., Latané et al., 1979). “Working together” in loafing experiments means that the outputs of the participants are pooled. That is, the outputs of the participants are summed, and the “group’s” performance is represented by this sum. The Latané et al. (1979) description suggests that the performances of the participants in this “loafing” condition are compared to the performances of participants working by themselves. However, in virtually all of these studies, the number of participants has been held constant. In the within-subjects loafing designs (e.g., Harkins et al., 1980; Ingham et al., 1974; Latané et al., 1979; Williams, Harkins, & Latané, 1981), a fixed number of participants performed individually and in various sized groups. Thus, in these experiments, the trials on which participants were asked to perform individually were like social facilitation trials in which participants performed sequentially rather than simultaneously.

In the between-subjects loafing designs (e.g., Brickner et al., 1986; Harkins & Jackson, 1985; Harkins & Petty, 1982), groups of participants, whose outputs were individually identifiable, performed the task simultaneously (e.g., vigilance, brainstorming) exactly as in facilitation research. Thus, in the prototypic loafing experiment, the performances of a set of participants whose outputs are pooled are compared to the performances of a set of participants whose outputs are individually identifiable.

The Latané et al. (1979) description of the loafing effect suggests that loafing is a “group versus individual” effect. However, our analysis suggests that loafing would be more appropriately viewed as an evaluation effect. In fact, Latané et al. (1979) suggested that the “information reducing” nature of the “group” trials (Davis, 1969) may have been responsible for the loafing effect. When participants’ performances were pooled, individual outputs were lost in the crowd, submerged in the total,
and were not individually recoverable by the experimenter. Because participants could receive neither praise nor blame for their performances, they loafed. Consistent with this interpretation, Williams et al. (1981) found that when participants were led to believe that their individual outputs could be monitored even when they performed "together" (i.e., pooled outputs), there was no loafing. Also, when participants were led to believe that interest centered on "group" performance and their individual outputs were to be summed, they loafed as much when they performed individually as when they performed "together." Williams et al. concluded on the basis of these findings that identifiability of individual outputs is an important mediator of social loafing.

However, Williams et al. did not manipulate identifiability alone. The participants all worked on the same task; thus, when their performances were individually identifiable, they could be directly compared to 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 identifiability alone, that motivated performance. To test this possibility, Harkins and Jackson (1985) orthogonally manipulated identifiability and comparability, using a brainstorming task in which participants were asked to generate as many uses as possible for an object. Replicating previous loafing research, when outputs were identifiable, participants generated more uses than when their outputs were pooled. However, this difference emerged only when participants believed that their outputs could be evaluated through comparison to their co-workers' performances.

These findings suggest that identifiability alone does not motivate performance. Participants must also feel that their outputs can be compared to those of others. Without this potential for evaluation, participants whose outputs were individually identifiable exerted as little effort as those whose outputs were pooled. These data suggest that people are motivated to work when their outputs can be evaluated. When outputs are pooled, evaluation is not possible and it is this aspect of "working together" that leads to loafing.

Viewed in this way, the experimental conditions that have been included in loafing and facilitation research fall into three cells of a 2 (Alone vs. Coaction) × 2 (Evaluation vs. No Evaluation) factorial design. Coaction in this context is defined as more than one person working on the same task during the same period of time. The manner in which the individual scores are to be treated is left unspecified (e.g., individual evaluation, summed). In this 2 × 2 design, facilitation studies are characterized by the comparison of the performances of coactors whose outputs can be evaluated (Coaction/Evaluation) to the performances of single participants whose outputs can also be evaluated (Alone/Evaluation). In virtually all loafing research, the performances of participants in a Coaction/Evaluation
cell have been compared to the performances of coacting participants whose outputs were pooled (Coaction/No Evaluation). The fourth cell, Alone/No Evaluation, has been run in no extant coaction study using either the facilitation or loafing paradigms. We propose that this combined design incorporates the processes that account for both facilitation and loafing effects, and addresses weaknesses present in each paradigm taken alone.

For example, Markus (1981) argued that those studies in which the mere presence hypothesis has not been confirmed “are studies in which there has not been a clear ‘alone’ condition as a comparison baseline” (p. 259). Consistent with the notion that experimenters have not been sufficiently aware of this problem, Bond and Titus (1983), in a review of social facilitation research, wrote: “In 96 of 241 studies, the experimenter was in the room with the ‘alone’ subject, and in 52 of these studies this ‘alone’ subject could see the experimenter!” (p. 271). Guerin (1986) reviewed 287 social facilitation studies and found only 11 published experiments that met his criteria for a test of the mere presence hypothesis. However, it could be argued that even Guerin’s criteria, which included removing the experimenter from the room, were not strict enough for a reasonable test. Markus (1978) noted: “In virtually all experiments with humans, the subject in the alone condition is not ‘phenomenologically’ alone even when the experimenter is physically removed and out of sight. That is, he is quite aware of the experimenter and knows that his performance is being recorded, presumably for some present or future evaluation” (p. 391). What difference does it make whether the experimenter is out of the room, if she or he has immediate access to a record of the participants’ performances?

In all but one of the 11 published studies that Guerin cited as representing the best tests of the mere presence hypothesis, the experimenter had access to the participants’ scores immediately after completion of the task. The sole exception was the Markus (1978) study in which the mere presence manipulation consisted of an attentive or inattentive audience which was present when the participants donned and doffed familiar and unfamiliar clothing. Apparently, there are no human coaction studies that incorporate appropriate tests of the mere presence hypothesis.

In the combined design we have described, the abilities of potential sources of evaluation, including the experimenter, to evaluate individual performances are minimized in the “No Evaluation” conditions. Thus, the addition of the conditions suggested by loafing research may offer a closer approximation to the conditions necessary for a test of the effects of mere presence on coaction, using tasks like those commonly employed in facilitation research.

In the loafing studies that we have described, the number of people present has been held constant. In only one loafing experiment has the
performance of a person actually working alone been compared to the performances of people working together (Kerr & Bruun, 1981), and in this experiment, the number of people present and the manipulation of evaluation were confounded. That is, in the alone condition, a single person's outputs could be evaluated by the experimenter through comparison of his/her performance with the performances of the preceding and subsequent participants, while in the together conditions, there were at least two participants whose outputs were pooled and, therefore, could not be evaluated. Did participants in the alone condition put out more effort than the “pooled” participants because they were working alone or because their outputs could be evaluated? The description of the loafing effect points to the former possibility (e.g., Latané et al., 1979); however, the latter possibility is supported by the findings of Williams et al. (1981) and Harkins and Jackson (1985). The unified design clarifies the relationship between number and evaluation by orthogonally manipulating these variables.

**EXPERIMENT 1**

To test this unified paradigm, the complete $2 \times 2$ design was run in two experiments. In the first experiment, participants, run alone or in pairs, were asked to generate as many uses for an object as they could. Crossed with the number manipulation, one-half of the participants were told that everyone in the experiment was generating uses for the same object (Evaluation), while the other half were told that each of the participants would be generating uses for a different object (No Evaluation).

To replicate previous social facilitation research, participants in the Coaction/Evaluation condition should put out greater effort than Alone/Evaluation participants. To replicate previous social loafing research, Coaction/Evaluation participants should exert greater effort than Coaction/No Evaluation participants. If mere presence has an effect in this paradigm, Coaction/No Evaluation participants should put out greater effort than participants in the Alone/No Evaluation condition. This pattern of results would yield two main effects, suggesting that both “evaluation” and “number” affect performance. If mere presence has minimal effects, performance in the Alone/No Evaluation condition should be equivalent to performance in the Coaction/No Evaluation condition, yielding an Evaluation $\times$ Number interaction.

In making the case that loafing and facilitation effects can be accommodated in the same framework, we argued that pooling the participants' outputs leads them to reduce their efforts because it minimizes the possibility of evaluation. To provide additional support for this notion, a Coaction/Pooled Output condition was added to the design of the experiment. To replicate the findings of previous loafing research, Coaction/Evaluation participants should put out greater effort than
Coaction/Pooled Output participants, who should perform at the same level as Coaction/No Evaluation participants.

Method

Subjects

The subjects were 96 males and females who participated in the experiment as a means of earning partial course credit. The basic design was a 2 (Alone vs. Coaction (Pair)) × 2 (Evaluation vs. No Evaluation) with 12 participants run in each of the single conditions and 12 pairs in each of the pair conditions. An additional 12 pairs were run in a loafing replication condition in which the pair’s outputs were pooled.

Procedure

Upon arrival, the singles or pairs were seated at a semi-circular table with partitions that prevented the pairs from seeing one another, and were given instruction sheets that informed them that we were studying the performance of individuals and groups on a task called brainstorming. They would be given the name of an object and their task would be to generate as many uses as they could for the object. They were also told that we were interested in the number of uses that could be generated for a range of objects. Since some objects were easy and some difficult to generate uses for, the number of uses they generated would be comparable only to the number generated by others working on the same object.

In the Evaluation condition, participants were told that in this experiment we were looking at the number of uses that could be generated for one particular object, and so, everyone would be generating uses for the same object. The participants were asked to take 1 of 20 small envelopes. To the inside of the envelopes were paperclipped slips on which was written the name of the object for which they were to generate uses. They were reminded that it did not matter which envelope they took since everyone was generating uses for the same object. In each envelope there were also a number of slips of paper on which the uses were to be written.

In the No Evaluation conditions each participant was told that she or he was the only participant who would be presented with this particular object. They were also asked to take one of the 20 envelopes, but were told that each envelope contained a slip on which the name of a different object had been written.

The experimenter then invited the participants to remove the clipped object slip and to read it. In all cases, the participants were asked to generate uses for a knife. They were then asked to dump the blank slips in their envelopes on the table in front of them, to take one slip, to fold it three times, and to slide it down the tube in front of them that extended into a box. The top of the box was then removed and the participants were shown either a single bin (Singles) or a pair of bins (Pairs) into which their use slips would fall.

The participants were asked to write one use per slip, to fold each slip three times, and to slide it down their tube. All of the participants were asked to generate as many uses as they could and to not be concerned about the quality of their reactions. The uses could be ordinary or unusual. They were to simply try to generate as many uses as possible. They were given 12 min to generate their uses, a length of time that previous research (Harkins & Petty, 1982) would suggest provided more than enough time for them to generate as many uses as they could. The experimenter then closed a sound attenuated partition for the 12-min use listing period. After this period, the participants were asked to respond to a set of ancillary measures, were debriefed, and then dismissed.

An additional 12 pairs were run in a loafing replication cell in which pairs were asked to generate uses for the same object but the slips of each pair were deposited in a common
STEPHEN G. HARKINS

bin. Thus, in this condition No Evaluation was accomplished through pooling. These participants also responded to the ancillary measures.

Results

The scores for each pair were averaged resulting in 12 observations per cell. The basic data were analyzed in 2 (Alone vs Coaction) x 2 (Evaluation vs. No Evaluation) ANOVAs and Dunnett’s test (Kirk, 1982) was used to make comparisons involving the loafing replication cell (Coaction/Pooled output).

Uses

Analysis of the primary dependent variable, the number of uses, revealed two main effects. Evaluation participants generated more uses (M = 20.66) than No Evaluation participants (M = 16.25), F(1, 44) = 16.6, p < .01, and Coactors generated more uses per person (M = 20.45) than Singles (M = 16.45), F(1, 44) = 13.6, p < .01. The interaction was not reliable (p > .20).

Replicating the social loafing effect, participants in the Coaction/Evaluation condition generated more uses (M = 22.58) than participants in the Coaction/Pooled Output condition (M = 19.00, p < .05). The number of uses generated by Coaction/Pooled Output participants (M = 19.00) was not reliably different from the number generated by participants in the Coaction/No Evaluation condition (M = 18.33, p > .20).

Ancillary Measures

Among the ancillary measures, participants were asked to rate the extent to which they believed that the experimenter could determine exactly how many uses they individually generated. Of course, all of the participants in the basic 2 x 2 design should have believed that the experimenter could tell exactly how many uses they individually generated. Analysis of these data revealed no reliable differences (ps > .20). However, in the Coaction/Pooled Output, or loafing condition, the participants should have felt that the experimenter was less able to determine exactly how many uses they individually generated than participants in either the Coaction/No Evaluation or the Coaction/Evaluation conditions. After all, in both of the latter two conditions the number of uses generated by each person could be counted. Consistent with this analysis, the Coaction/Pooled Output participants felt that the experimenter was less able to determine the number of uses generated by each person (M = 5.08 on an 11-point scale) than participants in the Coaction/Evaluation (M = 7.83) or Coaction/No Evaluation conditions (M = 7.58, ps < .05).

Participants were also asked to rate the extent to which they believed that their performance in this experiment could be directly compared to
the performances of others. Participants in the Evaluation conditions were told that all participants were generating uses for the same object. Thus, they would be expected to report that their performances could be compared to those of others to a greater extent than participants in the No Evaluation conditions, who were told that they had different objects for which to generate uses, and they did ($M_{\text{evaluation}} = 7.58$, $M_{\text{no evaluation}} = 5.58$), $F(1, 44) = 16.1, p < .05$. When asked about comparability of performance, Coaction/Pooled Output participants reported that they believed that their performances could be compared to the performances of others in the experiment to the same extent ($M = 5.91$) as participants in the Coaction/No Evaluation condition ($M = 6.00, p > .20$), and believed that their performances could be compared less than participants in the Coaction/Evaluation condition ($M = 7.67, p < .05$).

**Discussion**

Coaction/Pooled Output participants generated fewer uses than Coaction/Evaluation participants, a replication of the loafing effect. Consistent with the notion that the potential for evaluation is central to the loafing effect, Coaction/No Evaluation participants generated the same number of uses as the participants in the Coaction/Pooled Output conditions, and reliably fewer than those in the Coaction/Evaluation condition. The ancillary measures were also supportive of this analysis. Participants in the loafing replication condition could not be evaluated because their outputs were pooled. The different object manipulation was intended to achieve the same end, and, consistent with this expectation, participants in these two conditions reported that their outputs could not be evaluated to the same extent, both reliably less than participants in the Coaction/Evaluation condition.

In the basic $2 \times 2$ design, two main effects were obtained. With evaluation held constant, coactors outperformed singles. With number held constant, participants whose outputs could be evaluated outperformed participants whose outputs could not be. This pattern of results suggests that both the potential for evaluation and mere presence affected performance.

The logic underlying the same object/different object manipulation, that was used in Experiment 1, suggests that two conditions must be met for evaluation to be possible. The participants’ individual outputs must be identifiable, and there must be some standard against which these outputs can be compared. In the typical loafing paradigm, participants generate uses for the same object, but, because their outputs are pooled, individual efforts cannot be evaluated. When participants have different objects, even though their outputs are identifiable, there is no standard for comparison. Results for the ancillary measures in the present research
were consistent with this analysis. Also, the identical same object/different object manipulation was used by Harkins and Jackson (1985), who also found that participants whose outputs were identifiable, but not comparable (different objects), generated as few uses as participants who generated uses for the same object, but whose outputs were pooled.

However, Harkins and Petty (1982) have used a similar manipulation and reported a different outcome. Participants whose uses were to be pooled were told that they would either be generating uses for the same object or for different objects. In these pooled conditions, Harkins and Petty (1982) found that different-object participants generated more uses than same-object participants, the opposite of what was found by Harkins and Jackson (1985) and in the present research. Though these experiments are superficially similar, there were a number of methodological differences that could account for this discrepancy. For example, in the Harkins and Petty (1982) research, participants in the different-object condition were told “you alone are responsible for generating uses for the object that you will be given” (p. 1226), while participants in Harkins and Jackson (1985) and the current research were not. Participants in the Harkins and Jackson (1985) research and the present work were told that some of the objects were easy and some difficult to generate uses for, and, as a result, the number of uses they generated would be comparable only to the number generated by others working on the same object, while the participants in the Harkins and Petty (1982) were not.

**EXPERIMENT 2**

To assess the reliability of the findings of Experiment 1, the effects of number and evaluation potential were examined in a second experiment, using a different method of manipulating evaluation potential. Participants took part as singles or pairs in a vigilance task that required them to report seldomly occurring signals presented on a TV screen. Evaluation was manipulated by leading the participants to believe either that the computer that kept track of their performances was functioning properly (Evaluation) or had malfunctioned (No Evaluation).

In addition to using a different method of manipulating evaluation potential, this vigilance task is “optimizing” (requiring a criterion level of performance, Steiner, 1972), rather than “maximizing” (requiring as much effort as possible, Steiner, 1972) like the brainstorming task used in Experiment 1, so we can see whether these effects generalize to a different type of task. Finally, in Experiment 2 when the computer fails, there is ostensibly no record of performance, unlike in the brainstorming task where there was a product (the number of uses), even though comparison was not possible because all of the objects were supposedly different. Thus, this vigilance task goes even farther toward minimizing the possibility of evaluation.
Method

Subjects

The subjects (Ss) were 90 male and female undergraduates who participated as a means of earning partial course credit. The design was a 2 (Alone vs. Coactors) × 2 (Evaluation vs. No Evaluation) factorial with 15 participants run in each of the Single cells and 15 pairs run in each of the Coaction cells.

Procedure

As the experimenter (E) was taking the participants to the laboratory, all of them were told that this research was only in its initial stages and all of the bugs had yet to be worked out. Upon arrival, participants were seated at a semi-circular table with partitions that prevented the pairs from seeing one another. Participants then read the following instructions that were presented on a TV monitor:

We are interested in studying the performance of groups and individuals on vigilance tasks. The vigilance task requires you to watch for a dot to flash on a TV screen. When you see the dot you are to signal by pressing a button. You will be watching one-fourth of the TV screen. In this session, you will be watching the upper lefthand quadrant of the screen. We are interested in your performance on this task. Over the course of the experiment we will be keeping track of the number of signals you detect. Please try to detect as many signals as you can, while minimizing the number of times you falsely report the presence of a signal.

All participants then took part in a one-minute practice trial during which three signals were presented. The 12-in. TV screen was divided into 4 equal areas by computer-generated graphics, and the Ss were reminded that they were to watch the upper lefthand quadrant. The signals, which flashed for 1/30th of a second and were each composed of one graphics block, occurred at randomly determined locations in this block and during this practice trial were presented at 15-sec intervals. After the first signal, the display was interrupted by the message ‘Response Recording Error.’ The E then went through a slightly opened partition into the other half of the room, saying that the response recorder had been malfunctioning. The E then returned and asked the Ss to try again. In the Evaluation condition, the Ss responded to the two remaining signals and no error message appeared. The Ss were then presented individual hit and false alarm feedback for the two signals they had seen. They were also told that because the experiment was just beginning, the E was interested in the Ss’ perceptions regarding the task, and so, at the end of the experiment they would be given a questionnaire which would require them to rate their experience along several dimensions.

In the No Evaluation condition, response to the second signal resulted in the same error message, and the E told these Ss that the response recorder would be disabled, but even so, the last signal would be presented to which they should also respond. At the end of the practice session, the feedback showed ‘0’ hits and ‘0’ false alarms for the Ss. The E then told them that the response recorder was malfunctioning, so their performances could not be recorded, but the E would like for them to perform the task anyway. Because the experiment was just beginning, the E was interested in the Ss’ perceptions concerning the task, and at the end of the experiment, the Ss would be given a questionnaire which would require them to rate their experience along several dimensions. Only if they took part in the vigilance task would they be able to answer these questions accurately, and so, the Ss were to try as hard as they could.

All Ss were then told that if any other error messages appeared on the screen they
should go the E's office down the hall to report it. Otherwise the computer would inform them when the task was over at which time they would be asked some questions about the experiment.

The E then closed the room partition and left the room for the 14-min duration of the task. During the 14 min, 14 signals were presented. For each session, the location of the signals within the quadrant was randomly determined, but the timing of the signals across the 14 min was the same for all sessions. After the viewing task ended, the E returned, the Ss responded to the questionnaire, were debriefed, and then dismissed.

Results

The scores for each pair were averaged resulting in 15 observations in each of the cells comprising the 2 (Coactors vs. Alone) × 2 (Evaluation vs. No Evaluation). For analysis, the two types of errors (misses and false alarms) were summed because, consistent with previous research (Harkins & Petty, 1982), a preliminary analysis revealed that these measures were positively related.

Analysis of the error index (misses and false alarms) revealed two main effects. Evaluation participants made fewer errors (M = 2.40) than No Evaluation participants (M = 7.48), F(1, 56) = 35.5, p < .01, and Coactors made fewer errors (M = 3.95) than Singles, M = 5.93), F(1, 56) = 5.4, p < .05. The interaction was not reliable (p > .20).

Among the ancillary measures, participants were asked to rate the extent to which the E could tell exactly how well they individually performed. Participants in the Evaluation condition reported that the E could tell to a greater extent how they performed (M = 8.26 on an 11-point scale) than No Evaluation participants (M = 4.21), F(1, 56) = 23.12, p < .01.

Discussion

The findings from these experiments refine our understanding of both loafing and facilitation. In both experiments, using very different tasks, main effects were obtained for number and evaluation. Coactors outperformed participants working alone, not the reverse as was suggested by the description of the loafing effect (e.g., Latané et al., 1979). Consistent with the findings of previous loafing research (e.g., Harkins & Jackson, 1985; Latané et al., 1979; Williams et al., 1981), participants whose outputs could be evaluated outperformed participants whose outputs could not be.

The two main effects are consistent with the argument that both mere presence and evaluation played a role in motivating performance. Of course, it is probably not possible to eliminate all concerns about the possibility of evaluation when participants know that they are taking part in an experiment. However, by minimizing the apparent opportunities for evaluation by the experimenter, the present experiments approached this goal more closely than previous coaction experiments in which no
Social loafing and social facilitation attempt was made to eliminate this source of evaluation. Though mere presence and evaluation apprehension have often been pitted against each other as rival explanations for facilitation effects, Markus (1981) pointed out that both factors could contribute to facilitation effects. In these experiments, both did, and in an additive fashion.

This analysis suggests that there is no discrepancy between the findings of loafing and facilitation research. In fact, the findings from these paradigms are complementary. In social facilitation research, when participants coact, their outputs can be compared and they work harder than participants working alone. In social loafing research, when participants coact, their outputs cannot be evaluated, and they put out less effort than participants whose outputs can be compared. In both cases, evaluation is central. In social facilitation, working together enhances evaluation potential; in social loafing, working together reduces it.

The Latané et al. original definition of loafing could be taken to be generic. That is, any motivation loss in groups could be called "social loafing." For example, Kerr (1983) found evidence of reduced effort that he termed "free rider" and "sucker" effects when disjunctive and conjunctive scoring schemes were used in groups. These findings would appear to fall under the rubric "social loafing," because participants working together put out less effort than those working individually. However, in these experiments, the participants' individual outputs were always identifiable. Hence, it is quite unlikely that these findings could be attributed to the effects of evaluation. The current analysis argues in favor of a restricted definition of loafing that focuses on the fact that "working together" leads to reduced effort because there is no opportunity for evaluation.

Our analysis has been limited to the simple, well-learned tasks that have been used in most loafing research. On tasks like rope-pulling, shouting, pumping air, and vigilance, greater effort leads to better performance. However, when complex tasks are used in facilitation research (e.g., Martens & Landers, 1972; Seta, Paulus, & Schkade, 1976), greater effort leads to poorer performance. This suggests that if, instead of the simple tasks typically used in loafing research, a complex task were used, participants in the Coaction/No Evaluation (loafing) condition should perform better than participants in the Coaction/Evaluation condition, and this outcome is exactly what was found by Jackson and Williams (1985). Participants in a Coaction/Evaluation condition performed better on simple mazes than participants whose outputs were pooled (Coaction/No Evaluation), a loafing effect, but when the mazes were complex, the pattern was reversed. Participants whose outputs were pooled performed better than participants whose individual outputs could be compared. Thus, it appears that this approach can account for performance on both simple and complex tasks.
The potential for evaluation plays a central role in this account of loafing and facilitation effects, but evaluation by whom? In 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 good performance or to avoid blame for a bad one. Whenever the experimenter was unable to monitor individual outputs directly, performers sloughed off" (p. 464). However, the experimenter is only one of three potential sources of evaluation. When outputs are pooled, participants may also feel that they cannot evaluate their own output, nor can this output be evaluated by their fellow participants.

Social facilitation researchers have referred to each of these potential sources of evaluation in their accounts of coaction effects (e.g., coactor evaluation, Klinger, 1969; self-evaluation, Sanders, Baron, & Moore, 1978; experimenter evaluation, Seta, Paulus, & Schkade, 1976). Despite the fact that each of these sources has been incorporated into these accounts of coaction effects, there is no compelling evidence that these accounts have captured the motivational structure of the experiments they were meant to describe. For example, though Sanders et al. (1978) allude to self-evaluation in their interpretation, when the participant could evaluate himself or herself, the coactors could also evaluate the participant, and at this time, the evaluation potential of the experimenter was likely to have been particularly salient. In many of these facilitation experiments, it is not clear what the participants were told or inferred about evaluation potential. The haphazard manipulation of evaluation potential may account for Bond and Titus (1983) finding that evaluation potential had no systematic effects in social facilitation research.

Thus, the most that can be said at this point is that the potential for evaluation plays an important role in motivating performance in this "rudimentary social arrangement" (Cottrell, 1972). By using the minimal evaluation conditions suggested by loafing research as a starting point, the motivational structure of this "rudimentary social arrangement" can be examined systematically, leading to a better understanding of both loafing and facilitation effects.

In this analysis, we have focused on the role that evaluation plays in producing facilitation and loafing effects. We are not proposing 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, 1983). Rather, we are arguing that evaluation potential plays a central role in producing the reduction in effort that has been termed social loafing, and the enhancement in effort that has been termed social facilitation. 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, 1986), personal involvement (Brickner et al., 1986), 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 sources was minimized. In future work, it will be necessary to determine how these other factors interact with the potential for evaluation to motivate performance in these settings.

REFERENCES


