Social Facilitation: Using the Molecular to Inform the Molar

Allison E. Seitchik, Adam J. Brown, and Stephen G. Harkins

The Oxford Handbook of Social Influence
Edited by Stephen G. Harkins, Kipling D. Williams, and Jerry Burger

Subject: Psychology, Personality and Social Psychology, Neuropsychology
Online Publication Date: Sep 2016 DOI: 10.1093/oxfordhb/9780199859870.013.8

Abstract and Keywords

Research conducted for more than a century has shown that the presence of others improves performance on simple tasks and debilitates it on complex tasks, whether these others are audience members or coactors. In this chapter, we review theories offered to account for how two features of these others, their mere presence and/or the potential for evaluation they represent, produce these effects, and we conclude that we are no closer now to isolating the relevant process(es) than we were 100 years ago. We then consider the molecular task analysis proposed by Harkins (2006) as an approach to attacking this problem, followed by a review of the work supporting the mere effort account suggested by this analysis. Finally, we place the mere effort account in the larger context represented by the Threat-Induced Potentiation of Prepotent Responses model, which aims to account for the effect of threat on task performance.

Keywords: social facilitation, mere presence, evaluation apprehension, motivation, threat, task performance

A chapter on social facilitation would typically begin with the claim that Triplett (1898) published the first experiment on social facilitation. In fact, Triplett's work has often been described as the first experiment published in social psychology. However, given Stroebe's (2012) choleric analysis of this claim, let it suffice to say that there has been interest in the effect that the presence of others has on task performance for well over a century. These effects have been studied in two paradigms. In one, the audience paradigm, the others are present as observers or spectators watching the task performance of the participant. In the other paradigm, coaction, the others are present working independently on the task alongside the participant.

In this chapter, we briefly review the early work on social facilitation. We consider research testing the two features of the presence of others that were hypothesized to produce the performance effects, their mere presence and apprehension concerning their

evaluation potential. We review the theories offered to account for how mere presence and evaluation apprehension produce these effects, and we conclude that we are no closer now to isolating the relevant process or processes than we were 100 years ago. We consider the molecular task analysis proposed by Harkins (2006) as an approach to attacking this problem, followed by a review of the follow-up work done to test the mere effort account suggested by this analysis. Finally, we place the mere effort account in the larger context represented by the Threat-Induced Potentiation of Prepotent Responses model, which aims to account for the effect of threat on task performance.

A Look Back

Triplett's (1898) work has been interpreted as showing that the presence of others enhanced performance, as was much of the work that followed (e.g., Allport, 1920, 1924), leading Allport (1924) to coin the term *social facilitation*. However, it was also found that the presence of others sometimes debilitated performance, and in his 1935 review, Dashiell was unable to specify when which outcome would be produced. Matters had not progressed by the 1950s when Asch (1952) observed that: "The suspicion then arises that the proffered concepts are simply restatements of the quantitative results" (p. 67). Although this criticism was directed specifically at Allport's (1924) research, it actually summarized the whole line of work for its first 50 years.

Perhaps as a result, interest in social facilitation waned until 1965, when Zajonc's drive interpretation of social facilitation renewed interest in this area. In his extremely influential *Science* article, Zajonc surveyed past research on audience (e.g., Bergum & Lehr, 1963; Dashiell, 1930; Pessin, 1933; Travis, 1925) and coaction effects (e.g., Allport, 1920; Gates & Allee, 1933; Travis, 1928) and suggested that these findings could be organized by the simple generalization that the presence of others, as spectators or as coactors, enhances the emission of dominant responses. By dominant response, Zajonc (1965) meant the response that was most probable in a participant's task-relevant behavioral repertoire. In the early stages of learning to perform a task, Zajonc (1965) argued that this response is likely to be incorrect, but, with practice, correct answers become more probable; that is, they become dominant.

As Zajonc (1965) noted, Spence (1956) had already established that arousal, activation, or drive results in the enhancement of dominant responses. Zajonc (1965) then described indirect, but suggestive, evidence supporting the notion that the presence of others increases a person's arousal level. Thus, his resolution suggested that presence of others produces arousal that enhances the likelihood of the emission of the dominant response.

If the task is simple or well learned, the dominant response is likely to be correct, and, as a result, performance will be facilitated, but if the task is complex or the appropriate responses are not yet mastered, the dominant response is likely to be incorrect, and the performance will be debilitated.

This paper had a galvanizing effect on research in social facilitation. In a cumulative graph of research activity in social facilitation, Guerin (1993) showed that following the publication of Triplett's work in 1898, a small number of articles were published each year from 1920 through 1965, but, at this point, the number of articles shot up, and this rate of activity continued through 1983, the endpoint of the graph. Although other interpretations for facilitation effects were offered (e.g., Blank, Staff, & Shaver, 1976; Duval & Wicklund, 1972), most of the research focused on the drive account. Zajonc (1965) suggested that the mere presence of others was sufficient to increase drive, whereas others argued that it was the evaluation and/or competition associated with the others that produced the drive. For example, Cottrell (1968) argued: "If coaction and performance before an audience usually result in positive or negative outcomes for the individual, then he will quickly come to anticipate these outcomes when he coacts with others or performs before an audience" (p. 104), and Cottrell (1968) proposed that it was these associations that produced the drive effects. This proposal led to a series of studies in which researchers tried to design experiments to pit these explanations against each other (e.g., Cottrell, Wack, Sekerak, & Rittle, 1968; Henchy & Glass, 1968; Paulus & Murdoch, 1971).

Whether it was mere presence or evaluation apprehension that produced drive, the drive theory account of social facilitation remained the dominant theory of the time. For example, in 1977, Geen and Gange wrote: "A review of the literature on social facilitation following Zajonc's advocacy of drive theory as the explanatory principle shows that this theory, in general, provides the most parsimonious explanation of the findings reported over those 12 years" (p. 1283). However, after another 12 years, matters had changed. In his review of the facilitation literature since 1977, Geen (1989) wrote: "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" (p. 17). Geen (1989) proposed three classes of these alternative approaches.

One set of theories included those approaches that continued to rely on the proposal that the presence of others increased drive (e.g., distraction/conflict [Baron, 1986]; evaluation apprehension [Cottrell, 1972]; social monitoring [Guerin & Innes, 1982]; and compresence [Zajonc, 1980]). The second set was comprised of approaches that proposed "the presence of others creates either explicit or implicit demands on the person to

behave in some way" (Geen, 1989, p. 31) (e.g., self-presentation [Bond, 1982]; self-awareness [Carver & Scheier, 1981]). The third set argued that the presence of others affects the focus of attention and information processing (e.g., an information-processing view of distraction/conflict effects suggested as an alternative to his drive account by Baron [1986]). At this point, the controversy concerning mere presence and evaluation apprehension was still unresolved. Each of the theories that Geen (1989) reviewed incorporated mere presence effects, evaluation apprehension effects, or both.

Tests of Mere Presence

Separating the effects of mere presence and evaluation apprehension was recognized as a daunting challenge early on. For example, in his 1935 review, Dashiell noted, "to get pure 'alone' or pure 'co-working' (and may we add, pure 'spectator' or pure 'competing') situations is extraordinarily difficult" (p. 1115). This difficulty has persisted to the present day. In virtually all social facilitation experiments, the experimenter could represent a potent source of evaluation in the "alone" conditions. There has been some recognition of this problem in its most egregious form. For example, in a review of social facilitation research, Bond and Titus (1983) 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 (1993) raised this issue again in his review of mere presence effects. After excluding studies "if they clearly involved group discussion, imitation or the exchange of reinforcements" (p. 129), Guerin (1993) was left with 313 studies. After applying the 12 criteria that he judged appropriate for a good test of the mere presence hypothesis, including the removal of the experimenter from the room, he was left with 18 studies. Of these 18 studies, Guerin (1993) reported that "eleven found evidence for mere presence effects and seven did not" (p. 137). However, in fact, Guerin's (1993) criteria were not stringent enough. As 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). Guerin (1993) did not exclude experiments that afforded the experimenter the opportunity to evaluate performance in the alone condition at the conclusion of the session because "this might, in fact, remove all mere presence tests since all have used a laboratory task which subjects know will be evaluated by the experimenter" (p. 135). However, the reduction of all sources of evaluation to a minimum is exactly what is required in the "no evaluation" conditions. Of Guerin's (1993) published

studies, only four experiments in which performance was measured met this more stringent criterion. Two are audience experiments (Markus, 1978; Schmitt, Gilovich, Goore, & Joseph, 1986), each of which found evidence consistent with the mere presence hypothesis.

The other two are coaction experiments (Harkins, 1987) in which participants in “no evaluation” conditions were led to believe that their performances could not be evaluated by anyone during or after the experiment. In the 2 (alone versus coaction) x 2 (no-evaluation versus evaluation) design, mere presence was manipulated by asking participants to work alone, or alongside another person working independently on the same task (coaction). Crossed with this manipulation, one half of the participants were led to believe that their performances would be evaluated after their performance, whereas the other half were led to believe that their performances would not be evaluated.

Using two different, simple tasks (use-generation and vigilance), Harkins (1987) found main effects for both mere presence and evaluation in this design. Coactors outperformed participants working alone, a “mere presence” effect, and participants whose outputs could be evaluated outperformed participants whose outputs could not be, an evaluation effect. It is probably not feasible 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, Harkins’s (1987) research approached this goal more closely than previous coaction experiments in which the most that was done was to remove the experimenter from the room.

Subsequent to Guerin’s (1993) review, other attempts have been made to test the mere presence hypothesis. For example, Huguet, Galvaing, Monteil, and Dumas (1999) tested the attentional view of social facilitation effects (Baron, 1986) by examining the effect of social presence on the performance of the Stroop color-word task. They led participants to believe that the computer on which they would be performing the Stroop color-word task had not yet been programmed to record their responses. They were asked to perform the task anyway so that they could give their impressions about this new task at the end of the session. They performed the task alone or in the presence of different types of audience (inattentive, invisible, or attentive). Huguet et al. (1999) found that Stroop interference (that is, the difference in the time taken to identify the ink colors of incongruent words, such as the word “red” printed in green ink, and the ink color of control stimuli, for example, + + + printed in green) was reduced when the audience was invisible or attentive, but not when it was inattentive. Huguet et al. (1999) argued that this finding was consistent with Guerin’s (1986) argument that mere presence effects

emerge when there is uncertainty about the behavior of the audience (e.g., when the behavior of the audience members cannot be monitored).

However, Klauer, Herfordt, and Voss (2008) proposed that these effects were the result of the fact that to make sense of the fact that they were performing the task on an unprogrammed computer, participants were instructed to try to form a general impression of the task. According to Klauer et al., this instruction led to unusually long reaction times in the alone condition, which produced the interference effect. In response, Sharma, Booth, Brown, and Huguet (2010) noted that their follow-up research "refutes Klauer et al.'s (2008) position. Here the dual task instructions were removed, but the social facilitation remained, at least under long RSI (the time between the response and the following stimulus), as expected" (pp. 56–57, parenthetical phrase mine). Although this change may have eliminated this alternative interpretation, it also reintroduced the experimenter as a potential source of evaluation. That is, whether or not the experimenter was actually in the room, the participants could have believed that she or he would be able to evaluate the participant's performance after the session concluded. As a result, there is no true "alone" condition in these experiments. Thus, while this work (Klauer et al., 2008; Sharma et al., 2010) may show social presence effects, it does not provide an adequate test of mere presence.

Platania and Moran (2001) asked participants, who were either alone or observed by another student, to make verbal judgments on the relative size of stimulus squares as they were presented on a wall. The authors report that they found that the participants in the audience condition responded with their preferred response numbers (responses with the highest habit strength) more often than participants who were alone. They argued that the "socially facilitated dominant responses were personal choices (numerical preferences) that could not be construed as either right or wrong, and, therefore, were irrelevant to evaluative judgments" (p. 196). Thomas, Skitka, Christen, and Jurgena (2002) exposed participants who were either alone or in a group composed of a real participant and two confederates (coaction) to an experimenter who attempted to make a very favorable or very unfavorable impression on them. After completing a phrase-completion task, the ostensible point of the study, participants were asked to anonymously rate the experimenter as part of the departmental research process. The experimenter left the room, and the evaluations were placed into envelopes that the participants put in a box full of other envelopes. Consistent with the mere presence hypothesis, the participants in the coaction condition rated the experimenter who had attempted to make a favorable impression more favorably and the experimenter who had attempted to make a negative impression less favorably than participants who took part alone. In these experiments, the potential for evaluation is apparently eliminated, and, along with the experiments of Markus (1978) and Schmitt et al. (1986), they provide

evidence consistent with the mere presence hypothesis (see also Greenier, Devereaux, Hawkins, & Johnston, 2001).

Tests of Evaluation Effects

Taken together, the findings suggest that mere presence contributes to social facilitation effects. As noted previously, when Zajonc (1965) proposed his drive account, research efforts focused on whether the effects were produced by mere presence *or* evaluation apprehension. The research reviewed thus far suggests that mere presence is sufficient to produce social facilitation effects, but it does not suggest that mere presence is the only explanation. As pointed out by Markus (1981), both factors could contribute to facilitation effects.

In fact, this is the outcome reported in the Harkins (1987) research. In that research, participants worked alone or coacted, and their performances could be evaluated or not. For two tasks, Harkins (1987) found main effects for evaluation and presence. This research supports an additive model with mere presence and the potential for evaluation each contributing to the social facilitation effects on these simple tasks. And it provides evidence consistent with the mere presence hypothesis: Participants in the coaction/no-evaluation condition outperformed participants in the alone/no-evaluation condition. However, interpretation of the evaluation conditions is problematic in this research and in the great majority of the other research looking at the contribution of the potential for evaluation to social facilitation effects.

For example, in the alone/evaluation condition of the Harkins (1987) research, the experimenter had access to the participants' outputs, and so could evaluate them by comparing them to the performances of previous participants. Because this same opportunity for evaluation existed in the coaction/evaluation condition, it may seem that the improved performance in this condition must be the result of mere presence because the potential for evaluation has been held constant. But, in fact, the potential for evaluation was not held constant. In the alone/evaluation condition, the experimenter could evaluate participants only by comparing their performance to that of their predecessors, but, in the coaction/evaluation condition, there were at least three different ways in which the potential for evaluation could have been increased.

First, participants might feel more evaluation apprehension when they know that the experimenter can compare their performances not only to those of their predecessors but also to those of coactors who are present in the same session (the experimenter evaluation account). Second, participants could feel more apprehension in the coaction

setting when they know that they themselves can compare their performances to those of participants who are present in the same session (the self-evaluation account). Third, participants could feel more apprehension when they know that their performances can be evaluated by coactors who are present in the same session (the coactor account).

In Harkins's (1987) research, in the alone/evaluation and coaction/evaluation conditions, there was the potential for evaluation by the experimenter, but in the coaction condition, there was also the potential for coactor and self-evaluation. Latané's (1981) theory of social impact would suggest that when a person is the target of social forces emanating from other people, the magnitude of the effect of these sources *would* be a multiplicative function of the strength, immediacy, and number of people present. Thus, as the number of sources increases, the amount of impact should be increased, resulting in improved performance on a simple task. Of course, social impact theory would only make clear predictions in those cases in which there is a manipulation of the number of *external* sources (for example, increasing the number of coactors, or coactors plus experimenter). It would not make a prediction if one of the sources of evaluation were the self.

In fact, in a series of experiments (Bartis, Szymanski, & Harkins, 1988; Harkins, 2000, 2001a, 2001b; Harkins & Lowe, 2000; Harkins & Szymanski, 1988, 1989; Harkins, White, & Utman, 2000; Seitchik & Harkins, 2014; Szymanski & Harkins, 1987, 1993; Utman & Harkins, 2010; White, Kjelgaard, & Harkins, 1995), Harkins and his colleagues have found that, when taken alone, the potential for experimenter and self-evaluation each motivates performance, but when both are possible, concern over the potential for experimenter evaluation supersedes interest in self-evaluation. These findings suggest that the potential for self-evaluation may not contribute to facilitation effects, at least in combination with the potential for experimenter evaluation.

Of course, this research leaves open the possibility that the potential for self-evaluation can contribute to facilitation effects in combination with the potential for evaluation by a coactor. However, Szymanski, Garczynski, and Harkins (2000, Experiment 1) found that concern over the potential for evaluation by a coactor appears to supersede interest in the potential for self-evaluation in just the same way as the potential for evaluation by the experimenter. In a second experiment, Szymanski et al. tested the coactor account and found that the motivation stemming from the potential for evaluation by the coactor summed with the motivation produced by the potential for evaluation by the experimenter. Thus, consistent with Latané's (1981) theory of social impact, the combination of these sources (experimenter plus coactor) led to better performance than either one taken alone.

Gagné and Zuckerman (1999) also report findings consistent with this analysis. Their participants were led to believe that no one would be able to evaluate their performance on a use generation task, that the experimenter alone would be able to do so, or that the experimenter and coactors could evaluate them. The linear contrast on these data was significant, showing that the combination of the potential for experimenter and coactor evaluation produced better performance than the potential for experimenter evaluation alone, which produced better performance than when evaluation by no one was possible.

These experiments represent the beginning of what will be required to understand the effects of mere presence and the potential for evaluation in the social facilitation paradigm. In many experiments, researchers leave out conditions that would allow a systematic examination of the effects of evaluation on performance, because these conditions may not bear on the particular question that they are asking. For example, Gagné and Zuckerman (1999) did not include a condition in which there was the potential for coactor evaluation alone. Jackson and Williams (1985) included an alone/evaluation condition along with coaction/evaluation and coaction/no-evaluation conditions, but not an alone/no-evaluation condition. Sanna (1992) included an alone/no-evaluation condition along with coaction/evaluation and coaction/no-evaluation conditions, but no alone/ evaluation condition.

In addition, in research in which the potential for evaluation was manipulated, it is not clear what the participants were told or inferred about which source or sources would have access to their performances. The haphazard manipulation of this variable may account for Bond and Titus's (1983) failure to find that evaluation potential had systematic effects in their meta-analysis of social facilitation research. To tie down these evaluation effects will require a set of experiments that focus on each of the potential sources of evaluation in the facilitation paradigm individually (for example, experimenter, self, and coactor), and then in combination with the other sources, along with the appropriate tests of mere presence effects. It is only at this point that we will know how to design performance settings to capture the motivational structure that we want. For example, if we believe that behavior motivated by the potential for self-evaluation will be more likely to be sustained over the long haul than that produced by external evaluation, we may wish to structure situations such that the potential for external evaluation is minimized.

Mediator(s) of Social Facilitation Effects

However, even if we knew how mere presence and the potential for evaluation combine to produce social facilitation effects, we are still left with the task of specifying which, if any, of the theories incorporated in Geen's (1989) review can account for these effects. It is noteworthy that these theories do not even agree on whether participants subject to evaluation are working hard on complex tasks, but failing nonetheless (or as a consequence), are withdrawing effort leading to failure, or are failing as a result of a cognitive, rather than a motivational process.

Drive approaches suggest that participants are putting out more effort when they work in the presence of others, and this increased effort enhances performance on simple tasks (correct answer high in the habit hierarchy) but debilitates performance on complex tasks (correct answer low in the habit hierarchy). Bond's (1982) self-presentation approach suggests that participants are working hard in the presence of others, but concern over the embarrassment of potential failure causes cognitive interference. Baron's (1986) focus-of-attention interpretation would suggest that it is not a motivational effect but a cognitive one. Participants working on difficult tasks are working as hard as participants facing easy tasks, but the presence of others leads to a narrowing of attention that debilitates performance on complex tasks because these tasks require attention to a wide range of cues.

In contrast to these approaches, Carver and Scheier (1981) suggested that debilitated performance on complex tasks results from the withdrawal of effort. That is, the presence of others makes participants self-aware, leading them to be "more cognizant of both the level of performance being manifested at the moment and the salient standard" (Geen, 1989, p. 32). When they find that they have little chance of bringing their performance into alignment with the standard, they stop trying.

Twelve years after Geen's review, in their own review, Aiello and Douthitt (2001) described what are essentially the same theories and presented a framework for advancing theory in this area. Unfortunately their advice was not followed. In fact, in looking back over the work done in this area, it appears that, instead of making any progress, the same set of potential explanations are just recycled. For example, Triplett (1898) suggested "brain worry" as one explanation for his findings, which reappeared as "cognitive interference" in the 1980s (e.g., Bond, 1982) and is now termed "working memory deficit" (e.g., DeCaro, Thomas, Albert, & Beilock, 2011). Another explanation, range of cue utilization (Easterbrook, 1959), which was used by Geen (1976) and Baron

(1986) to account for performance effects, remains with us but is now termed “focus of attention” (e.g., Muller & Butera, 2007).

One could argue that the mediating process will emerge from an analysis of the psychophysiology of performance. For example, Blascovich has applied his biopsychological model of challenge and threat to the phenomenon of social facilitation (Blascovich, Mendes, Hunter, & Salomon, 1999). Blascovich et al. (1999) found the pattern of cardiovascular reactivity associated with challenge when participants performed a well learned task in the presence of others, whereas on a task that was not well learned (that is, complex), they exhibited the pattern of cardiovascular reactivity associated with threat. In contrast, participants working alone demonstrated no reactivity from baseline, whether the task was well learned or not.

This work identifies precise physiological patterns of cardiovascular reactivity that are associated with the facilitation and debilitation of performance. However, as Blascovich et al. (1999) note, it does not tell us whether the physiological responses are causes, concomitants, or results of the performance effects. In addition, because the timing period was limited to the first minute of the testing period, we know only that the participants performing the unlearned task in front of the audience were displaying the cardiac pattern associated with threat at that point. We do not know whether these participants continued to exhibit this patterning as the timing period continued, nor do we know whether these participants continued their striving in an effort to perform the task successfully or if they withdrew effort. As a result, we are left with the same questions that are left unresolved in the other accounts.

Mere Effort Account

Harkins (2006) argued that our lack of progress might be a result of the fact that our efforts have focused broadly on theory construction rather than on an analysis of how performance actually unfolds on a given task. Although it would appear that a molecular analysis of task performance would be an integral part of theory development, this type of analysis has not been conducted, and it is possible that the mediating process could be identified through such an approach. To this end, Harkins (2006) undertook a molecular analysis of the effects of evaluation on the performance of a specific task, the Remote Associates Test (RAT).

The RAT requires participants to look at a set of three words (e.g., *lapse, elephant*, and *vivid*) and generate a fourth word that is related to each word in the given triad (in this

case *memory*). Harkins (2001b) has shown that the potential for evaluation produces the typical pattern of performance on this task: Participants who anticipate evaluation by the experimenter solve more triads shown by a pretest to be simple than do no-evaluation participants, whereas participants who anticipate experimenter evaluation solve fewer triads shown by a pretest to be difficult than do no-evaluation participants.

Harkins's (2006) analysis suggested the *mere effort* account. This explanation argues that the potential for evaluation motivates participants to want to do well, which potentiates whatever response is prepotent ("dominant," using Zajonc's [1965] terminology) on the given task. For example, when performing the Stroop Color-Word Test (Stroop, 1935), the prepotent response is to read the color-word, rather than call out the color, which the mere effort account argues is potentiated by the potential for evaluation. On the RAT, Harkins's (2006) analysis showed that the prepotent response is to generate words that are closely related to one of the triad members. Because on simple items the correct answers tend to be a close associate of at least one of the triad members, the greater effort on the part of participants subject to evaluation leads to the production of more close associates and to better performance.

On the other hand, on the complex items, the associations between the triad members and the correct answer are much weaker (i.e., the associates are more remote), and the participants are extremely unlikely to produce the solution by generating associates for the individual triad members. For example, if presented with the triad member *note*, a participant would be extremely unlikely to produce the associate, "bank." Nonetheless, when the participant considers the word *note*, the solution, "bank," is weakly activated. Likewise, the solution "bank" is also weakly activated when the participant considers the other two triad members, *river* and *blood*. If this were the only process operating, this weak activation should accumulate over time, leading to the emergence of the correct answer. However, when participants actively test close associates as solutions for the triads (e.g., "teller" for "bank"), these associates are highly activated, and they strongly inhibit the activation of the remote (weak) associates. Thus, generating close associates, the same behavior that facilitates the performance of participants subject to experimenter evaluation on simple items, debilitates that performance on complex items.

Drive theory (Zajonc, 1965), like the mere effort account, accords a central role to prepotent or dominant responses. Once again, drive theory contends that the presence of others produces arousal, which increases drive, enhancing the probability of the emission of dominant responses. On simple tasks, these responses are likely to be correct; on difficult tasks they are likely to be incorrect. In fact, Cottrell (1968, 1972) argued that this drive was the result of the participants' apprehension about the fact that they could be evaluated.

Thus, both mere effort and Cottrell's (1968, 1972) evaluation apprehension account of social facilitation effects predict that the potential for evaluation will potentiate dominant or prepotent responses. However, in the case of mere effort, this potentiation results from the motivation to perform well, which should also lead to an effort to correct an incorrect response, if the participant recognizes that his or her response is incorrect, knows the correct response, and has the opportunity to make it. In contrast, Cottrell's (1968) modification of Zajonc's (1965) drive theory suggests only that the positive or negative anticipations produced by the presence of others nonselectively energize individual performance (i.e., potentiate the dominant response). On a task like the RAT, one is unable to distinguish between mere effort and evaluation apprehension accounts because even if the participants know that the response is incorrect, they do not know how to correct it. As a result, one cannot see the effect of the motivation to correct on this task.

However, these findings are inconsistent with the other three explanations. For example, Harkins's (2006) findings suggest that on the complex RAT items, participants subject to evaluation do not perform poorly because they withdraw effort (Carver & Scheier, 1981). It is the fact that they are putting out effort that is the source of their difficulty on complex triads. It is also not that worry concerning failure takes up processing capacity (Bond, 1982). Once again, Harkins's (2006) findings suggest that participants subject to experimenter evaluation are engaged in the same behavior on both simple and complex items. It is just that this behavior is effective on simple items but is ineffective on complex ones.

A third explanation, focus of attention, suggests that the potential for evaluation produces an attentional overload that "leads to a restriction in cognitive focus in which the individual attends more to cues that are most central to the task (or alternatively most central geographically in the display) at the expense of more peripheral cues" (Baron, 1986, p. 27). This cognitive explanation does not account for the role that motivation plays in producing the pattern of results on the RAT. That is, participants who are subject to evaluation do not perform better on simple items because the answer candidates that they generate are more closely related to the triad members (central cues) than are the answer candidates generated by the no-evaluation participants. The participants in the two conditions are equally likely to think of answer candidates that are closely related to the triad members. It is the fact that participants subject to evaluation are motivated to *generate and test* more of these closely related candidates that accounts for the fact that they outperform no-evaluation participants. On the complex items, it is this same motivation to test more closely related candidates that inhibits the accumulation of the activation required for the correct answer to emerge. Thus, no-evaluation participants do not perform better on complex items than do participants who

are subject to evaluation because no-evaluation participants are better able to think of more remotely associated answer candidates (i.e., peripheral cues) than are participants subject to evaluation. No-evaluation participants perform better because the same lack of motivation that prevents them from testing enough closely related candidates to come up with correct answers on simple items allows the small amount of activation produced by each triad member to accumulate to the point that the correct answer *pops out* on the complex items.

Although Harkins's (2006) findings are inconsistent with these three accounts, this research was aimed at attempting to identify the mediating process, rather than at pitting this process against these accounts. In addition, this work does not distinguish between the drive/evaluation apprehension and mere effort accounts. McFall, Jamieson, and Harkins (2009) conducted a set of experiments that were aimed at pitting the mere effort account against these other explanations, and, in each case, they found support for the mere effort account.

For example, McFall et al. (2009) used the Stroop Color-Word task to test the focus of attention account as well as to distinguish between the mere effort and drive/evaluation apprehension explanations. Consistent with Baron's (1986) analysis, Huguet et al. (1999; see also Huguet, Dumas, & Monteil, 2004) have found that social presence enhances performance on the Stroop and have argued that this facilitation is a result of the fact that social presence reduces the range of cues used by the participants. As they write, "Narrowing one's focus should indeed allow one to screen out the incorrect semantic cues and focus more exclusively on the letter color cues" (Huguet et al., 1999, p. 1013). That is, these participants see less of the word, and so it interferes less with their response.

In their review of previous work on the Stroop, Huguet et al. (1999) cited work that they suggested shows that "arousal has been associated with *increased* (italics added) Stroop interference in past research" (p. 1012; e.g., Hochman, 1967, 1969; Pallak, Pittman, Heller, & Munson, 1975). They cited other research that shows that "distraction has been associated with *decreased* (italics added) Stroop interference in past research" (p. 1013; e.g., Houston, 1969; Houston & Jones, 1967; O'Malley & Poplawsky, 1971), as well as MacKinnon, Geiselman, and Woodward's (1985) research, which shows that coaction decreases Stroop interference. Huguet et al. (1999) commented on the contradictory nature of these findings and argued that their well-controlled experiments show that social presence reduces the amount of Stroop interference, consistent with the focus of attention interpretation.

However, McFall et al. (2009) argued that the findings of the previous work are not contradictory, and, in fact, are consistent with the mere effort account. Consistent with the mere effort and drive accounts, reading the word is the prepotent (dominant) response. However, on the Stroop, unlike the RAT, the fact that this response is incorrect is quite obvious, as is the correct response. Thus, when given sufficient time, participants who are more motivated can inhibit the prepotent response and still produce the correct response more quickly than can participants who are less motivated.

In the previous research that showed increased Stroop interference, the participants had only 1 second to produce the response (Hochman, 1967, 1969; Pallack et al., 1975), and the dependent measure was the number of errors. Under these conditions, the more motivated participants did not have sufficient time to inhibit the prepotent response to read the word and make the correct response. Thus, the more motivated participants made more errors than did less motivated participants. In the experiments that showed decreased Stroop interference (Houston, 1969; Houston & Jones, 1967; Huguet et al., 1999; MacKinnon et al., 1985; O'Malley & Poplawsky, 1971), the dependent measure was the time required to read a whole list of color-words or the time required to make each individual response. Under these conditions, the motivated participants had sufficient time to inhibit the reading response and still produce the color response more quickly than did the less motivated participants.

This analysis suggests that simply manipulating the amount of time available for the response can produce either facilitation or debilitation. When given a limited response window (e.g., 1 second), participants subject to evaluation should make reliably more errors than participants who are not. That is, the prepotent response is to read the word, and given a brief response period, participants subject to evaluation will not have enough time to inhibit this response and generate the correct response. When more time is made available for the response (e.g., 2 seconds), few mistakes should be made and participants subject to evaluation should name the colors more quickly than participants who are not subject to evaluation.

In contrast, the focus of attention account predicts better performance by participants subject to evaluation at each response window. Reducing the amount of time available for a response should not diminish the advantage afforded by a restricted focus of attention. In fact, if anything, one could argue that the restricted focus would lead to a greater performance advantage at the brief exposure period because these participants only see, and only need to see, part of the color word to respond correctly.

Finally, the drive/evaluation apprehension account (Cottrell, 1972; Zajonc, 1965) would predict that the presence of others simply increases drive, energizing the dominant response, reading the word. Thus, this account would predict that participants subject to

evaluation would perform more poorly than participants in the no-evaluation condition, regardless of the time available for a response.

McFall et al. (2009) tested these predictions in two experiments. They found that when the response window was brief (1 second or 750 ms; Experiment 3), participants subject to evaluation performed more poorly than no-evaluation participants. However, with the 2-second response window (Experiment 2), the finding was reversed: Participants subject to evaluation performed better than no-evaluation participants.

Finding that participants subject to evaluation make more errors on color-words than do no-evaluation participants is consistent with mere effort and drive/evaluation apprehension predictions. Each of these accounts would contend that reading color-words is the prepotent (dominant), but incorrect, response and that the motivation produced by the potential for evaluation potentiates this response. However, the drive/evaluation apprehension account only predicts this energization, whereas the mere effort account also predicts that participants subject to evaluation will be motivated to produce the correct answer. At the brief response windows (Experiment 3), these participants do not have sufficient time to inhibit the prepotent response and produce the correct response, and, as a result, they perform more poorly than no experimenter evaluation participants. However, at the longer response window (Experiment 2), they are able to inhibit the incorrect response and produce the correct response faster than no experimenter evaluation participants.

Finding that participants subject to evaluation perform better than do participants in the no-evaluation condition at the 2-second window is consistent with the focus of attention prediction. However, finding that they perform worse when brief response windows are used is not. If anything, a briefer display period should represent an advantage for participants with a narrowed focus of attention. Instead, we find that participants subject to evaluation make more errors than do participants who are not. Thus, in this test of the focus of attention and mere effort explanations, we find support for the mere effort account, but not focus of attention.

Recently Augustinova and Ferrand (2012) reported other findings that are inconsistent with Huguet et al.'s (1999) focus of attention account. In addition to the standard Stroop items (e.g., the word "blue" printed in green), these investigators included "color-associated" items (e.g., the word "sky" printed in green) in a replication of Huguet et al.'s (1999) research and found interference effects on both types of items in both their social presence condition and their alone condition. However, they also found that, unlike on standard Stroop items, social presence did not affect the magnitude of this effect for "color-associated" items. Of course, the "color-associated" items could only produce an interference effect if the participants were processing the semantic content of the words,

suggesting that the participants, regardless of social presence condition, are reading the words, not focusing “exclusively on the letter color cues” (Huguet et al., 1999, p. 1013).

Augustinova and Ferrand (2012) argue that this finding suggests that “social presence simply influences response competition” (p. 1216), not the semantic processing of the color-words. Augustinova and Ferrand speculate about how inhibitory processes suggested by Sharma et al. (2010) could work under these circumstances, but neither their findings nor the findings of Sharma et al. are inconsistent with the potentiation process proposed by the mere effort account.

However, Augustinova and Ferrand’s (2012) finding that social presence did not produce facilitation on either control or “color-associated” items is inconsistent with the mere effort account. Mere effort would argue that the social presence effect that Huguet et al. (1999) and Augustinova and Ferrand found on standard Stroop items stems from the fact that these participants are motivated to perform well and have the time required to correct for the prepotent response (reading the word) as in Experiment 2 of McFall et al. (2009). Reflecting this motivation to perform well, in Experiment 2, McFall et al. also found that participants subject to evaluation responded more quickly to the Stroop control stimuli than no-evaluation participants. They also found this facilitation effect on the prosaccade task (Hallett, 1978), a control task used with another inhibition task, the antisaccade task (McFall et al., Experiment 4), and this facilitation effect on control stimuli has been found in other research in this line of work (e.g., Stroop: Jamieson & Harkins, 2011; antisaccade task: Jamieson & Harkins, 2007, 2011; Jamieson, Harkins, & Williams, 2010). Of course, in some cases, floor effects could account for the failure to find effects for control stimuli, but this would be much less likely to occur for stimuli that produce interference effects, like those found by Augustinova and Ferrand on the “color-associated” items. Additional research will be required to resolve this issue.

Recently, building on previous work by Muller and Butera (2007), Normand, Bouquet, and Croizet (2014) tested an elaborated version of the focus of attention account. They propose that when participants are under evaluative pressure, they implement a stronger attentional set than when they are not, which allows them to filter out stimuli that are not critical for task performance. Thus, a distractor that does not match the attentional set produced by the task will lead to better performance by participants subject to evaluative pressure than those that are not, whereas a distractor that does match the attentional set will lead to poorer performance by these participants.

To test this account, Normand et al. (2014) first replicated an experiment by Muller and Butera (Exp. 5, 2007) with minor modifications. Participants were asked to look at a fixation point that was displayed for one of four durations ranging from 850 to 1,500 ms.

A dot then flashed for 30 ms approximately 7° from the screen's center in one of four locations forming a square followed by a 50 ms blank, and a display of four letters (three Qs and one O) in the same square. The participants' task was to identify the location of the O. On one third of the trials, the dot flashed in the location where the O would appear. On one third of the trials, the dot flashed in a location where a Q would appear, and on one third, dots flashed in all four locations. In other words, one third of the time the dot was a valid cue as to the location of the O, one third of the time the dot was an invalid cue, and one third of the time the four dots comprised control cues. Ordinarily, one would expect that a cue with a sudden onset would capture participants' attention (e.g., Yantis & Jonides, 1990), but according to Normand et al., the dot does not match the task demands (identifying the location of the letter O) and is, therefore, irrelevant. As a result, participants subject to evaluative pressure should be able to filter out its effects, making them faster on the trials when the cue is invalid but slower on the trials when the cue is valid than participants not subject to this pressure.

Consistent with this prediction, Normand et al. (2014) reported that they replicated Muller and Butera's (2007) finding: a smaller cueing effect (i.e., the difference between the reaction times for invalid and valid cues) for participants subject to evaluative pressure than for those that were not. This finding stands in contrast to the prediction that they ascribe to the mere effort account. The mere effort account argues, "evaluative pressure increases the emission of the 'prepotent' response, that is the response that is dominant for a given task" (p. 2). "In accordance with this account, research indicates that individuals under evaluative pressure are more prone to attentional automatisms, for example, they show higher propensity to reflexive saccades in the antisaccade task" (p. 2). This logic would suggest that participants who are subject to evaluative pressure should look at the abrupt onset cue more, not less, than participants who are not, and, as a result, the cueing effect should be larger, not smaller.

Normand et al.'s (2014) finding that participants subject to evaluation exhibit a smaller cueing effect than those who were not would appear to favor their focus of attention account over the mere effort account. However, this conclusion does not take into account the fact that the mere effort account incorporates correction. That is, as McFall et al. (2009) showed, when participants subject to evaluation know what the correct answer is (as on the Stroop), and are given adequate time, they outperform no-evaluation participants. Clearly on Normand et al.'s task, participants know the difference between the correct (O) and the incorrect (Qs) answers. In addition, Normand et al. gave participants as much time as they required to make their responses. As a result, participants subject to evaluation could look at the abrupt onset cue more than no-evaluation participants and still have faster reaction times than these participants.

As a result, finding a smaller cueing effect in reaction times alone does not allow us to distinguish between these two accounts. The most straightforward test would simply require the replication of Normand et al.'s (2014) experiment with eye tracking, which would allow us to determine whether participants subject to evaluative pressure look at the cue more or less than participants who are not. As a first step in this test, we have collected data from a baseline sample of 21 undergraduate participants (i.e., experimenter evaluation was not manipulated) in Normand et al.'s paradigm, while also collecting eye-tracking data.

The participants in this baseline sample showed the typical cueing effect in their reaction time (RT) data, such that invalid RTs were longer than valid RTs. More important, participants looked at the cue on only 24.6 percent of the invalid trials. We would argue that this percentage represents our best estimate of the potency of the cue. Participants responded to valid "cues" on 49.8 percent of the trials, but on these trials we cannot distinguish between eye movements to the cue and to the target, because they occupy the same space. Consistent with the argument that eye movements on valid trials consist of these two different responses, we find that the launch times for eye movements to invalid cues are faster than movements to valid cues. As a result, in this paradigm, our best estimate is that the cue attracts an eye movement approximately 25 percent of the time. Clearly this is not a prepotent response, and, as a result, we cannot distinguish between these two accounts using these stimulus parameters. Furthermore, the eye-tracking data revealed that the letters in the stimulus display were so large that participants did not need to make an eye movement at all to respond correctly. In fact, participants did not make eye movements on 30 percent of all trials in which they correctly identified the location of the letter.

More generally, these findings emphasize the importance of understanding the characteristics of one's task prior to using it to examine the effect of evaluation, or any other social variable, on performance. Clearly, using a terminal measure, like reaction time, is not sufficient to draw any conclusions about the processes that produce any differences that are found. These results indicate that the stimulus conditions used by Normand et al. (2014), as well as Muller and Butera (2007, Experiment 5), while producing an effect of evaluation on overall RTs, cannot be used to pit the focus of attention and mere effort accounts against each other. It will be necessary to replicate Normand et al.'s visual cueing experiment with eye tracking to determine what process is producing their evaluation effect. Going forward, the size of the stimuli could be reduced such that eye movement is required, but it would be necessary to demonstrate in a baseline experiment that looking at the cue is the prepotent response in this paradigm. For example, in this focus of attention work, the cues are displayed for very brief periods of time (e.g., 30 ms), and this short exposure time may affect the potency of the cue.

Normand et al. use other tasks in this set of experiments that provide evidence consistent with the focus of attention account, but, once again, the evidence consists only of reaction times. Without evidence about the underlying processes, we are unable to draw any firm conclusions about the source of these effects.

However, we can pit the focus of attention account against mere effort by using a visual attention task that has a well-established prepotent response and also requires that participants make eye movements to perform the task. The antisaccade task (Hallett, 1978) serves this purpose well. On the antisaccade task, a participant is asked to fixate on a cross that appears in the center of the visual display and to respond to a target presented randomly on one side of the display or the other. However, before the target appears, a cue (a white square) is presented on the opposite side of the display. Participants are instructed to ignore this cue and look to the opposite of the display where the target will appear. However, there is a reflexive-like, prepotent tendency to look at the cue that must be inhibited to optimize performance. Thus, this task shares many features with the visual cueing task used by Muller and Butera (2007) and Normand et al. (2014). For instance, both tasks begin with the presentation of a central fixation, followed by an abrupt onset peripheral cue, and then the target (central cue). Also, attention to the peripheral cue is not necessary for target identification in either task, and each task requires that participants shift their visual attention to the target's location to respond accurately. However, unlike the visual cueing task, on the antisaccade task, the peripheral cue is always on the side opposite to the one on which the target will appear.

Jamieson and Harkins (2007) used this task to test a mere effort account of the effect of stereotype threat on performance. Stereotype threat, like the potential for evaluation, arouses participants' concern about their ability to perform well on tasks relevant to the stereotype. Thus, the mere effort account argues that stereotype threat should produce the same basic pattern of findings on the antisaccade task as is produced by the potential for evaluation on other inhibition tasks like the Stroop. When not given sufficient time to correct for the prepotent tendency (i.e., at a brief display time), the participants subject to stereotype threat should be less accurate than controls in their ability to correctly identify target orientation. However, when the display time is increased enough to allow enough time for correction, stereotype threat participants should be able to respond to the target more quickly than controls, as a result of increased motivation to perform well, and this is exactly what Jamieson and Harkins (2007) found.

Jamieson and Harkins (2007) used an eye tracker to conduct a more fine-grained analysis of performance on this antisaccade task at a display time that permitted correction. Under these conditions, the mere effort account predicted that the participants under threat would look the wrong direction, toward the cue, more often than would

participants in the control group, because the motivation to perform well potentiates the prepotent response. At this point, if the participants have failed to inhibit the reflexive saccade, their eyes are at the cue and they must launch a corrective saccade to get to the target site. If they have successfully inhibited the saccade, they must launch a correct saccade to the target site from the fixation point. Because correct and corrective saccades are each an "extreme example of a voluntary saccade" (Sereno, 1992, p. 92), the motivation to correct should reduce the latency to launch each type of saccade, and, as a result, the evaluation participants should launch these saccades faster than control participants. Finally, after the participants' eyes arrive at the target area, the participant must determine the target's orientation and press the appropriate response key. When the participants see the target, the mere effort account predicts that the greater motivation of participants subject to stereotype threat would lead them to respond more quickly than would participants in the control condition. Jamieson and Harkins (2007) found support for each of these predictions.

McFall et al. (2009, Experiment 4) replicated this experiment with an evaluation manipulation. Overall, evaluation participants reported target orientation more quickly than no-evaluation participants with no sacrifice in accuracy. Of course, these findings are consistent with the predictions of both the focus of attention and the mere effort accounts. The accounts can only be pitted against each other when the processes that culminate in the terminal performance measures are examined. In each case, the eye-tracking measures are consistent with the mere effort predictions but not the focus of attention predictions. Participants subject to evaluation made more, not fewer, reflexive saccades. However, they also launched correct and corrective saccades faster than no-evaluation participants, as well as produced faster adjusted reaction times (times adjusted for the time of arrival of the participants' eyes at the target site).

As a result, even though evaluation participants looked the wrong way more often than no-evaluation participants, evaluation participants ended up outperforming the no-evaluation participants. These findings do not support the focus of attention account. Instead of focusing on the central cue (the target), the participants subject to evaluation looked toward the peripheral cue (the box) more, not less, than no-evaluation participants. The focus of attention account also cannot account for the motivated behavior reflected in the faster saccade launches for correct and corrective saccades, nor can it account for the faster adjusted reaction times. These findings are consistent with the mere effort account and replicate the pattern of findings that Jamieson and Harkins (2007) report in their research on stereotype threat.

It should be noted that the mediating process that Muller and Butera (2007) and Normand et al. (2014) invoked could lead to a different set of predictions for performance on the antisaccade task. That is, the focus of attention account argued that self-evaluation

threat leads participants to ruminate about the discrepancy between their performance and the participants' standards, which takes up attentional capacity, leading to a restricted focus of attention. Similarly, Schmader and Johns (2003) have argued that when under stereotype threat, participants expend cognitive resources that could be devoted to task performance on processing information resulting from the activation of the negative stereotype. Thus, in each case, participants are using processing capacity to ruminate about their task performance. However, instead of leading to reduced focus of attention, Schmader and Johns argued that the reduction in working memory capacity directly produces the performance debilitation reported in the stereotype threat literature (e.g., Cadinu, Maass, Rosabianca, & Kiesner, 2005; Croizet et al., 2004).

More specifically, Schmader and Johns (2003) argued that the executive attention component (central executive) of working memory (Engle, 2001, 2002) is impaired by the ruminations. The central executive is essential for effective performance on inhibition tasks, like the antisaccade task. Thus, if evaluation potential produces ruminations, which interfere with working memory, participants subject to evaluation should produce more reflexive saccades than controls because participants subject to evaluation have less ability to inhibit their tendency to look at the cue. They should also launch correct and corrective saccades more slowly than control participants because the capacity to launch these eye movements also requires the central executive (Kane, Bleckley, Conway, & Engle, 2001; Roberts, Hager, & Heron, 1994; Stuyven, Van der Goten, Vandierendonck, Claeys, & Crevits, 2000; Unsworth, Schrock, & Engle, 2004). However, as described previously, McFall et al. (2009, Experiment 4) found that participants subject to evaluation generated more reflexive saccades than control participants; they also launched correct and corrective saccades faster than control participants, and had faster adjusted reaction times. Each of these effects indicates that participants' central executive processes were not impaired by the potential for evaluation. Thus, these findings are consistent with the mere effort account but with neither the focus of attention account nor with Schmader and Johns's (2003) working memory account.

Taken together, McFall et al.'s (2009) research supports the mere effort account over these alternative explanations for social facilitation effects (withdrawal of effort, focus of attention, processing interference [aka working memory deficit]). In three other research traditions in psychology, it has also been found that the potential for evaluation tends to facilitate performance on simple tasks but to debilitate it on complex ones: social loafing (e.g., Jackson & Williams, 1985), creativity (e.g., Amabile, 1979), and achievement goal theory (e.g., Elliott, Shell, Henry, & Maier, 2005). Within these traditions, process models have been proposed to account for these findings, but a review reveals no agreement across, or even within, these traditions (Harkins, 2001b). In fact, the explanations that have been proposed to account for the performance effects are the same ones suggested

to account for social facilitation effects: concern about failure leads to withdrawal of effort (achievement goal theory [Elliott et al., 2005]; social loafing [Harkins, 2001b]; creativity [Hennessey, 2001]); concern over failure diminishes processing capacity (achievement goal theory [Elliott et al., 2005]); and attentional overload restricts focus of attention leading to poor performance on complex tasks, which often require use of a wider range of cues than simple tasks (creativity [Hennessey, 2001]). Thus, to the extent that mere effort can provide a plausible account for social facilitation effects, this account is likely to also apply to these other areas of research.

As noted earlier, Jamieson and Harkins (2007, 2009, 2011) have extended the analysis to the effects of stereotype threat on performance. Of course, when one is subject to stereotype threat, there is not only the possibility that one will fail to measure up as an individual, as is the case for the potential for evaluation, but also the possibility that one will confirm the negative stereotype, which is also hypothesized to motivate participants to perform well. As noted previously, Jamieson and Harkins (2007) found support for the mere effort account on the antisaccade task, and they have also found support in research using GRE quantitative problems (Jamieson & Harkins, 2009; Seitchik, Jamieson, & Harkins, 2014) and the Stroop task (Jamieson & Harkins, 2011). Additional research has found support for this account using horizontal subtraction problems, modular arithmetic (Seitchik & Harkins, 2015), and a virtual ball-bouncing task (Huber, Seitchik, Brown, Sternad, & Harkins, 2015).

Threat-Induced Potentiation of Prepotent Response Model

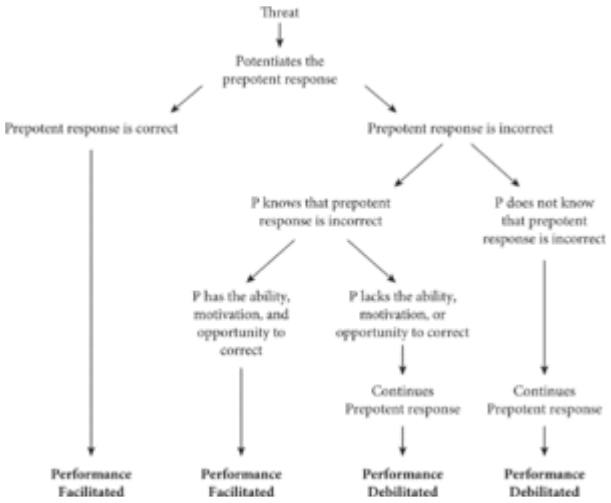
The research conducted thus far suggests that mere effort can account for performance on a range of tasks across domains. However, we would argue that these effects are best understood as part of a more encompassing account, the Threat-Induced Potentiation of Prepotent Responses (TIPPR) model. The TIPPR model builds on the mere effort account, arguing that the potential for evaluation and stereotype threat impacts task performance by potentiating prepotent responses, but they do so because they represent social *threats*, and it is the threat that produces the potentiation.

Consistent with this analysis, research suggests that exposure to social threats like stereotype threat and the potential for evaluation can produce a pattern of physiological arousal that is characterized by the activation of one or both of the two primary stress systems: the HPA (hypothalamus, pituitary, adrenal) axis; and the SAM (sympathetic adrenal medullary) axis. For example, in a meta-analysis of studies examining cortisol

levels in response to acute stressors, Dickerson and Kemeny (2004) found that exposure to social-evaluative threat was associated with heightened levels of cortisol, which is released in response to HPA activation. Rohleder, Wolf, Maldonado, and Kirschbaum (2006) found that a psychosocial stressor increased levels of salivary alpha amylase (sAA), a protein found in saliva that has been used as a proxy for catecholamines (specifically epinephrine and norepinephrine), which is released as part of the activation of the SAM axis. Blascovich and Mendes (2010) have summarized work in this area as showing that social threat, among other variables (e.g., effort and distress, striving for control, uncertainty, fear), is associated with moderate to high levels of activity in the HPA and SAM axes.

Furthermore, increases in arousal have been argued to be associated with the potentiation of prepotent (dominant) responses in both the evaluation (e.g., Cottrell, 1968, 1972) and stereotype threat (e.g., Ben-Zeev, Fein, & Inzlicht, 2005; O'Brien & Crandall, 2003) literatures. However, the TIPPR model argues that not just social threat but any environmental event that is perceived as a threat potentiates prepotent responses, and research on nonsocial threats is consistent with this argument. For example, Grillon, Ameli, Woods, Merikangas, and Davis (1991) have shown that the startle reflex produced by a loud auditory stimulus is potentiated by the threat of shock. Lang and his colleagues (e.g., Lang & Bradley, 2008) have shown that the viewing of unpleasant (threatening) stimuli also potentiates the startle reflex. Valls-Solé and his colleagues (e.g., Queralt et al., 2008) have shown that auditory startle speeds up obstacle avoidance. In this case, the researchers argue that the participants have prepared the motor program necessary for avoidance (i.e., it is prepotent), and the auditory startle (the independent variable in this case) potentiates this behavior. Similar effects have been found for walking, gait initiation, and sit-to-stand.

The potentiation of responses like these could have been adaptive in our ancestral past because responses to threat would likely require "flight or fight" or some other relatively simple behavior that would be facilitated by such potentiation. However, given the range of tasks that now confront us, this potentiation may help or hurt performance, as depicted in Figure 1.



Click to view larger

Figure 1 Summary of findings for research on the Threat-Induced Potentiation of Prepotent Responses (TIPPR) model.

If the prepotent response is correct, performance is facilitated. Thus, it is prepotency, not the fact that the task is "easy," that determines whether or not performance is facilitated. As it happens, on most "easy" tasks, the prepotent response is correct. However, one can decouple prepotency and task difficulty and see the primacy of prepotency.

For example, the GRE-quantitative test includes two types of problems: *solve problems*, which require the application of a formula; and *comparison problems*, which require logic and estimation. Previous work shows that participants' prepotent tendency when given math problems is to take a conventional or solving approach (i.e., compute the answer using a rule or an equation) (Gallagher & De Lisi, 1994; Gallagher et al., 2000; Jamieson & Harkins, 2009; Quinn & Spencer, 2001). Using these types of problems, Jamieson (2009) found the same pattern of performance whether GRE-Q items were more (50 percent average solution rate in the population) or less difficult (75 percent average solution rate). In each case, when the prepotent response (using the solving approach) is correct, participants subject to stereotype threat performed better than no-threat participants.

Likewise, McFall et al. (2009, Experiment 1) found that participants subject to evaluation outperformed no-evaluation participants on an anagram task when the words began with consonants (prepotent response is to try consonants as the first letter; Witte & Freund, 2001), whether the words to be solved were low frequency (difficult) or high frequency (easy). Thus, regardless of task difficulty, if the prepotent response was correct, performance on the task was facilitated. On other tasks, like the Remote Associates Task (RAT), however, task difficulty is confounded with prepotency. Harkins (2006) found that on the "easy" RAT items, the prepotent response (producing close associates of the triad members) was correct, and its potentiation facilitated performance.

If the prepotent response is incorrect, the outcome depends on whether or not the participant knows that this is the case. In many cases, the participant will not know that the prepotent response is incorrect. Under these circumstances, once again, it is

prepotency, not task difficulty that determines the outcome. For example, on comparison-type GRE-Q problems, the prepotent response (the conventional, or solving approach) tends to be incorrect, and threat participants perform more poorly than control participants whether the problems are harder (50 percent solution rate) or easier (75 percent solution rate). On anagrams, the prepotent response, trying consonants in the first position of the word, is incorrect on words that begin with vowels, and participants anticipating evaluation perform more poorly than no-evaluation controls, whether these words are low frequency (harder) or high frequency (easier). On the RAT, on "difficult" items, participants do not know that producing close associates of triad members, the prepotent response, is actually reducing the likelihood that activation of the correct answer will accumulate to the point that the correct answer "pops out" (Harkins, 2006), and potentiation debilitates performance.

Although, on many tasks, participants will not know that the prepotent response is incorrect, on some tasks, they do know, and performance on these tasks is informative as to the underlying processes. For example, on inhibition tasks, like the Stroop and the antisaccade task, the prepotent response is incorrect, but it is quite obvious to the participant that this is the case. Under these circumstances, performance depends upon whether the participant has the ability, motivation, and opportunity to correct for the incorrect, prepotent tendency.

On inhibition tasks like the Stroop and the antisaccade task, participants certainly have the ability to make the response, and the TIPPR model argues that the potential for evaluation and stereotype threat motivate participants to perform well. Given ability and motivation, the model predicts facilitation if participants have the opportunity to make the correct response, and this is what we find on the Stroop for evaluation and stereotype threat (McFall et al., 2009, Experiment 2; Jamieson & Harkins, 2011, respectively) and on the antisaccade task for evaluation and stereotype threat (McFall et al., 2009, Experiment 4; Jamieson & Harkins, 2007, Experiments 2 and 3, respectively). In each of these cases, even though the incorrect prepotent response is potentiated, participants subject to evaluation or stereotype threat outperform control participants.

We have also found that if participants are motivated to perform well, have the requisite ability, and the opportunity to make the response, simple instructions are enough to make up for the fact that participants do not know that the prepotent response is incorrect. For example, Harkins (2006) told all of his participants that their performance on the RAT would be subject to evaluation. One third of these participants were told that if they wanted to succeed, they should refrain from generating close associates. Instead they were to simply register the triad members, and then wait for the answer to "pop up." Another third were told that if they wanted to succeed, they should generate as many close associates as possible. The final third, a control condition, were told nothing.

Harkins (2006) found that participants who were told not to generate close associates, but to wait for the answer to emerge, outperformed the other two groups, which did not differ from each other.

Jamieson and Harkins (2009, Experiment 2) attempted to improve the performance of stereotype threat participants on the GRE-Q by providing instructions designed to reduce their reliance on the solving approach on comparison problems. This manipulation had a significant impact on performance, as the difference between threat and no-threat participants in the percentage of comparison problems solved was eliminated. A mediation analysis showed that threat participants' improvement on comparison problems resulted from the fact that threatened females took the instructions into account and reduced their reliance on the solving approach when they tried to solve comparison problems (see also Seitchik et al., 2014).

Even if participants know that the prepotent response is incorrect, if they lack the ability, motivation, or the opportunity to correct for the prepotent response, performance will be debilitated. For example, we have found that even when the participants are motivated and have the ability to make the response, if the time to respond is too brief (no opportunity to correct for the potentiation of the prepotent response), performance is debilitated on the Stroop (McFall et al., 2009, Experiment 3) and on the antisaccade task (Jamieson & Harkins, 2007, Experiment 1).

We have also found that performance is debilitated if threatened participants are not motivated to correct for the potentiated response, even if they have the requisite ability and the opportunity for correction. In the case of stereotype threat and evaluation, the effects of threat and motivation to do well cannot be separated. For instance, there is no stereotype threat without the concern that one's performance may confirm the stereotype. However, this is not the case for another social threat, ostracism. If one was ostracized, and then performed the antisaccade task, the experience of ostracism represents a threat. However, ostracized participants have no reason to be motivated to perform well unless their performance could serve to fortify their fundamental needs (belonging, self-esteem, control, and meaningful existence; Williams, 2009). In the absence of instructions suggesting this possibility, there is no reason for them to believe that the antisaccade task would serve this purpose. Under these conditions, the TIPPR model argues that the threat resulting from the experience of ostracism should potentiate the prepotent response, looking the wrong way at the cue. However, ostracism should have no effect on the measures of motivated behavior (volitional saccade launch times and reaction time adjusted for the time of arrival of the eyes at the target site) because the participant's needs cannot be fortified. This prediction was confirmed (Jamieson et al., 2010, Experiment 1) and shows that performance was debilitated if participants

lacked the motivation to perform well, even though they had both the ability and the opportunity to make the correction.

In a second experiment, an evaluation manipulation either directly linked performance on the antisaccade task to the experience of ostracism or did not (Jamieson et al., 2010, Experiment 2). When there was no link, the first experiment's findings were replicated (performance debilitated as a result of potentiated prepotent response: looking the wrong way at the cue), but when ostracized participants believed that their performance would be seen by their ostracizers, they were motivated to perform well on the antisaccade task, as reflected in faster volitional saccade latencies and adjusted reaction times, resulting in faster terminal reaction times, even though the ostracized participants also produced more reflexive saccades, the prepotent response. Once again, in this case, the participants had the ability to make the response, they had the opportunity, and they were motivated to do so.

Summary

Research relevant to understanding the effect that the presence of others has on task performance has been conducted for over a century (Stroebe, 2012). Taken together, this research has shown that mere presence and the potential for evaluation improve performance on simple tasks and debilitate it on complex tasks, whether the others are audience members or coactors. However, there is no agreement as to what process or processes mediate(s) these effects. In fact, the same explanations that were suggested in the early twentieth century are still with us today (e.g., focus of attention, working memory deficits, withdrawal of effort, drive theory).

Harkins (2006) suggested that a molecular analysis of performance on a single task might reveal a mediating process that could then be tested on other tasks. This analysis has led to the development of the Threat-Induced Potentiation of Prepotent Reponses (TIPPR) model, which is aimed at accounting for the effect of threat on task performance. This model builds on Harkins and his colleagues' (Harkins, 2006; McFall et al., 2009) mere effort account of social facilitation effects, which was then extended to stereotype threat (Huber et al., 2015; Jamieson & Harkins, 2007, 2009, 2011; Seitchik et al., 2014; Seitchik & Harkins, 2015). Other work (e.g., Jamieson et al., 2010) suggests that the mere effort account is best understood in the broader context provided by the TIPPR model.

The TIPPR model argues that threat potentiates prepotent responses. This potentiation could have been adaptive in our ancestral past because responses to threat would likely require "flight or fight" or some other relatively simple behavior that would be facilitated

by such potentiation. However, given the range of tasks that now confront us, this potentiation may help or hurt performance. If the prepotent response is correct, the model predicts that threat will facilitate performance. If the prepotent response is incorrect, and participants do not know, or lack the ability, motivation, or opportunity for correction, performance will be debilitated. However, if participants are able to recognize that their prepotent tendencies are incorrect and have the ability, motivation, and opportunity required for correction, performance will be facilitated. In previous work, Harkins and his colleagues have found support for this model using three different manipulations of social threat (the potential for evaluation: Harkins, 2006; McFall et al., 2009; stereotype threat: Jamieson & Harkins, 2007, 2009, 2011; and ostracism: Jamieson et al., 2010), and a variety of tasks (anagrams, antisaccade task, Remote Associates Task, GRE-quantitative problems, Stroop, horizontal subtraction problems, modular arithmetic, a virtual ball-bouncing task).

In 1935, Dashiell noted the difficulty of conducting research in this domain and then commented: “But this is no counsel of despair: Before solid findings can be achieved in any science a great amount of grub-work must needs be done in the way of clarification of problems and trying-out of techniques” (p. 1115). It is clear that much grub work remains both with respect to work that has already been done (e.g., accounting for the inconsistency with Augustinova and Ferrand’s (2012) findings; see also Wuhr and Huestegge, 2010) as well as to extensions of the account. However, even with the many issues that remain to be tackled, we believe that this approach has promise. The TIPPR model has relevance for any line of research that focuses on motivated task performance, such as social facilitation, social loafing, goal setting, intrinsic motivation/creativity, achievement goal theory, and stereotype threat, to name a few. It has the potential to integrate research across these areas by suggesting a common process through which threat affects performance. In addition, the work suggests that this process can be located within a more general model of reactions to both social and physical threat.

Acknowledgments

This work was funded in part by the U.S. Army Research Institute for the Behavioral and Social Sciences (Contract W5J9CQ-12-C-0046; PI: Stephen G. Harkins). The views, opinions, and/or findings contained in this report are those of the authors and shall not be construed as an official Department of the Army position, policy, or decision, unless so designated by other documents.

References

Aiello, J., & Douthitt, E. (2001). Social facilitation from Triplett to electronic performance monitoring. *Group Dynamics: Theory, Research and Practice*, *5*, 163–180. doi: 10.1037/1089-2699.5.3.163

Allport, F. H. (1920). The influence of the group upon association and thought. *Journal of Experimental Psychology*, *3*, 159–182. doi:10.1037/h0067891

Allport, F. H. (1924). *Social psychology*. New York, NY: Houghton Mifflin.

Amabile, T. M. (1979). Effects of external evaluation on artistic creativity. *Journal of Personality and Social Psychology*, *37*, 221–233. doi:10.1037/0022-3514.37.2.221

Asch, S. (1952). *Social psychology*. Englewood Cliffs, NJ: Prentice-Hall, Inc. doi: 10.1037/10025-000

Augustinova, M., & Ferrand, L. (2012). The influence of mere social presence on Stroop interference: New evidence from the semantically-based Stroop task. *Journal of Experimental Social Psychology*, *48*, 1213–1216. doi:10.1016/j.jesp.2012.04.014

Baron, R. (1986). Distraction-conflict theory: Progress and problems. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. *19*, pp. 1–40). New York, NY: Academic Press. doi:10.1016/S0065-2601(08)60211-7

Bartis, S., Szymanski, K., & Harkins, S. G. (1988). Evaluation and performance: A two-edged knife. *Personality and Social Psychology Bulletin*, *14*, 242–251. doi: 10.1177/0146167288142003

Ben-Zeev, T., Fein, S., & Inzlicht, M. (2005). Arousal and stereotype threat. *Journal of Experimental Social Psychology*, *41*, 174–181. doi:10.1016/j.jesp.2003.11.007

Bergum, B. O., & Lehr, D. J. (1963). Effects of authoritarianism on vigilance performance. *Journal of Applied Psychology*, *47*, 75–77. doi:10.1037/h0043188

Blank, T. O., Staff, I., & Shaver, P. (1976). Social facilitation of word associations: Further questions. *Journal of Personality and Social Psychology*, *34*, 725–733. doi: 10.1037/0022-3514.34.4.725

Blascovich, J., & Mendes, W. B. (2010). Social psychophysiology and embodiment. In S. T. Fiske, D. T. Gilbert, & G. Lindzey (Eds.), *Handbook of social psychology* (5th ed., Vol. *1*, pp. 194–227). Hoboken, NJ: John Wiley & Sons Inc.

Blascovich, J., Mendes, W. B., Hunter, S., & Salomon, D. (1999). Social “facilitation” as challenge and threat. *Journal of Personality and Social Psychology*, *77*, 68–77. doi: 10.1037/0022-3514.77.1.68

Bond, C. F. (1982). Social facilitation: A self-presentational view. *Journal of Personality and Social Psychology*, *42*, 1042–1050. doi:10.1037/0022-3514.42.6.1042

Bond, C. F., & Titus, L. J. (1983). Social facilitation: A meta-analysis of 241 studies. *Psychological Bulletin*, *94*, 265–292. doi:10.1037/0033-2909.94.2.265

Cadinu, M., Maass, A., Rosabianca, A., & Kiesner, J. (2005). Why do women underperform under stereotype threat? Evidence for the role of negative thinking. *Psychological Science*, *16*, 572–578. doi:10.1111/j.0956-7976.2005.01577.x

Carver, C., & Scheier, M. (1981). The self-attention-induced feedback loop and social facilitation. *Journal of Experimental Social Psychology*, *17*, 545–568. doi: 10.1016/0022-1031(81)90039-1

Cottrell, N. B. (1968). Performance in the presence of other human beings: Mere presence, audience, and affiliation effects. In E. C. Simmell, R. A. Hoppe, & G. A. Milton (Eds.), *Social facilitation and imitative behavior* (pp. 91–110). Boston, MA: Allyn & Bacon.

Cottrell, N. B. (1972). Social facilitation. In C. G. McClintock (Ed.), *Experimental social psychology* (pp. 185–236). New York, NY: Holt, Rinehart & Winston.

Cottrell, N. B., Wack, D. L., Sekerak, G. J., & Rittle, R. H. (1968). Social facilitation of dominant responses by the presence of an audience and the mere presence of others. *Journal of Personality and Social Psychology*, *9*, 245–250. doi:10.1037/h0025902

Croizet, J. C., Desprès, G., Gauzins, M. E., Huguet, P., Leyens, J. P. & Mèot, A. (2004). Stereotype threat undermines intellectual performance by triggering a disruptive mental load. *Personality and Social Psychology Bulletin*, *30*, 721–731. doi: 10.1177/0146167204263961

Dashiell, J. F. (1930). An experimental analysis of some group effects. *Journal of Abnormal and Social Psychology*, *25*, 190–199. doi:10.1037/h0075144

Dashiell, J. F. (1935). Experimental studies of the influence of social situations on the behavior of individual human adults. In C. Murchison (Ed.), *A handbook of social psychology* (pp. 1097–1158). Worcester, MA: Clark University Press.

DeCaro, M. S., Thomas, R. D., Albert, N. B., & Beilock, S. L. (2011). Choking under pressure: Multiple routes to skill failure. *Journal of Experimental Psychology: General*, *140*, 390–406. doi:10.1037/a0023466

Dickerson, S. S., & Kemeny, M. E. (2004). Acute stressors and cortisol responses: A theoretical integration and synthesis of laboratory research. *Psychological Bulletin*, *130*, 355–391. doi:10.1037/0033-2909.130.3.355

Duval, S., & Wicklund, R. A. (1972). *A theory of objective self-awareness*. Oxford, England: Academic Press.

Easterbrook, J. A. (1959). The effect of emotion on cue utilization and the organization of behavior. *Psychological Review*, *66*, 183–201. doi:10.1037/h0047707

Elliott, A. J., Shell, M. M., Henry, K. B., & Maier, M. A. (2005). Achievement goals, performance contingencies, and performance attainment: An experimental test. *Journal of Educational Psychology*, *97*, 630–640. doi:10.1037/0022-0663.97.4.630

Engle, R. W. (2001). What is working memory capacity? In H. L. Roediger, J. S. Nairne, I. Neath, & A. M. Suprenant (Eds.), *The nature of remembering: Essays in honor of Robert G. Crowder* (pp. 297–314). Washington, DC: American Psychological Association. doi: 10.1037/10394-016

Engle, R. W. (2002). Working memory capacity as executive attention. *Current Directions in Psychological Science*, *11*, 19–23. doi:10.1111/1467-8721.00160

Gagné, M., & Zuckerman, M. (1999). Performance and learning goal orientations as moderators of social loafing and social facilitation. *Small Group Research*, *30*, 524–541. doi:10.1177/104649649903000502

Gallagher, A. M., & De Lisi, R. (1994). Gender differences in Scholastic Aptitude Test: Mathematics problem solving among high-ability students. *Journal of Educational Psychology*, *86*, 204–211. doi:10.1037/0022-0663.86.2.204

Gallagher, A. M., De Lisi, R., Holst, P. C., McGillicuddy-De Lisi, A. V., Morely, M., & Cahalan, C. (2000). Gender differences in advanced mathematical problem solving. *Journal of Experimental Child Psychology*, *75*, 165–190. doi:10.1006/jecp.1999.2532

Gates, M. F., & Allee, W. C. (1933). Conditioned behavior of isolated and grouped cockroaches on a simple maze. *Journal of Comparative Psychology*, *15*, 331–358. doi: 10.1037/h0073695

Geen, R. G. (1976). Test anxiety, observation, and range of cue utilization. *British Journal of Clinical and Social Psychology*, *15*, 253–259. doi:10.1111/j.2044-8260.1976.tb00032.x

Geen, R. G. (1989). Alternative conceptions of social facilitation. In P. B. Paulus (Ed.), *Psychology of group influence* (2nd ed., pp. 15–51). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Geen, R. G., & Gange, J. J. (1977). Drive theory of social facilitation: Twelve years of theory and research. *Psychological Bulletin*, *84*, 1267–1288. doi: 10.1037/0033-2909.84.6.1267

Greenier, K. D., Devereaux, R. S., Hawkins, K. C., & Johnston, M. D. (2001). Social facilitation: The quest for true mere presence. *Journal of Social Behavior and Personality*, *16*, 19–31.

Grillon, C., Ameli, R., Woods, S. W., Merikangas, K., & Davis, M. (1991). Fear-potentiated startle in humans: Effects of anticipatory anxiety on the acoustic blink reflex. *Psychophysiology*, *28*, 588–595. doi:10.1111/j.1469-8986.1991.tb01999.x

Guerin, B. (1986). Mere presence effects in humans: A review. *Journal of Experimental Social Psychology*, *22*, 38–77. doi:10.1016/0022-1031(86)90040-5

Guerin, B. (1993). *Social facilitation*. New York, NY: Cambridge University Press. doi: 10.1017/CBO9780511628214

Guerin, B., & Innes, J. M. (1982). Social facilitation and social monitoring: A new look at Zajonc's mere presence hypothesis. *British Journal of Social Psychology*, *21*, 7–18. doi: 10.1111/j.2044-8309.1982.tb00506.x

Hallett, P. E. (1978). Primary and secondary saccades to goals defined by instructions. *Vision Research*, *18*, 1279–1296. doi:10.1016/0042-6989(78)90218-3

Harkins, S. G. (1987). Social loafing and social facilitation. *Journal of Experimental Social Psychology*, *23*, 1–18. doi:10.1016/0022-1031(87)90022-9

Harkins, S. G. (2000). The potency of the potential for experimenter and self-evaluation in motivating vigilance performance. *Basic and Applied Social Psychology*, *22*, 277–289. doi: 10.1207/15324830051035983

Harkins, S. G. (2001a). The role of task complexity, and sources and criteria of evaluation in motivating task performance. In S. G. Harkins (Ed.), *Multiple perspectives on the*

effects of evaluation on performance: Toward an integration (pp. 99–131). Norwell, MA: Kluwer Press.

Harkins, S. G. (2001b). The three-variable model: From Occam's razor to the black box. In S. G. Harkins (Ed.), *Multiple perspectives on the effects of evaluation on performance: Toward an integration* (pp. 207–259). Norwell, MA: Kluwer Press.

Harkins, S. G. (2006). Mere effort as the mediator of the evaluation-performance relationship. *Journal of Personality and Social Psychology*, *91*, 436–455. doi: 10.1037/0022-3514.91.3.436

Harkins, S. G., & Lowe, M. D. (2000). The effects of self-set goals on task performance. *Journal of Applied Social Psychology*. *30*, 1–40. doi:10.1111/j.1559-1816.2000.tb02303.x

Harkins, S. G., & Szymanski, K. (1988). Social loafing and self-evaluation with an objective standard. *Journal of Experimental Social Psychology*, *24*, 354–365. doi: 10.1016/0022-1031(88)90025-X

Harkins, S. G., & Szymanski, K. (1989). Social loafing and group evaluation. *Journal of Personality and Social Psychology*, *56*, 934–941. doi:10.1037/0022-3514.56.6.934

Harkins, S. G., White, P. H., & Utman, C. H. (2000). The role of internal and external sources of evaluation in motivating task performance. *Personality and Social Psychology Bulletin*, *26*, 100–117. doi:10.1177/0146167200261010

Henchy, T., & Glass, D. C. (1968). Evaluation apprehension and the social facilitation of dominant and subordinate responses. *Journal of Personality and Social Psychology*, *10*, 446–454. doi:10.1037/h0026814

Hennessey, B. A. (2001). The social psychology of creativity: Effects of evaluation on intrinsic motivation and creativity of performance. In S. G. Harkins (Ed.), *Multiple perspectives on the effects of evaluation on performance: Toward an integration* (pp. 47–75). Norwell, MA: Kluwer Press.

Hochman, S. H. (1967). The effect of stress on Stroop color-word performance. *Psychonomic Science*, *9*, 475–476.

Hochman, S. H. (1969). Stress and response competition in children's color-word performance. *Perceptual and Motor Skills*, *28*, 115–118. doi:10.2466/pms.1969.28.1.115

Houston, B. K. (1969). Noise, task difficulty, and Stroop color-word performance. *Journal of Experimental Psychology*, *82*, 403–404. doi:10.1037/h0028118

Houston, B. K., & Jones, T. M. (1967). Distraction and Stroop color-word performance. *Journal of Experimental Psychology*, *74*, 54–56. doi:10.1037/h0024492

Huber, M., Seitchik, A. E., Brown, A., Sternad, D., & Harkins, S. G. (2015). A mere effort account of the effect of stereotype threat on performance of a racket task. *Journal of Experimental Psychology: Human Perception and Performance*, *41*, 525–541. doi:10.1037/xhp0000039

Huguet, P., Dumas, F., & Monteil, J. M. (2004). Competing for a desired reward in the Stroop task: When attentional control is unconscious but effective versus conscious but ineffective. *Canadian Journal of Experimental Psychology*, *58*, 153–167. doi:10.1037/h0087441

Huguet, P., Galvaing, M. P., Monteil, J. M., & Dumas, F. (1999). Social presence effects in the Stroop task: Further evidence for an attentional view of social facilitation. *Journal of Personality and Social Psychology*, *77*, 1011–1025. doi:10.1037/0022-3514.77.5.1011

Jackson, J. M., & Williams, K. D. (1985). Social loafing on difficult tasks: Working collectively can improve performance. *Journal of Personality and Social Psychology*, *49*, 937–942. doi:10.1037/0022-3514.49.4.937

Jamieson, J. P. (2009). The role of motivation in blatant stereotype threat, subtle stereotype threat, and stereotype priming. Psychology Dissertations. Paper 10. Retrieved from

Jamieson, J. P., & Harkins, S. G. (2007). Mere effort and stereotype threat performance effects. *Journal of Personality and Social Psychology*, *93*, 544–564. doi:10.1037/0022-3514.93.4.544

Jamieson, J. P., & Harkins, S. G. (2009). The effects of stereotype threat on the solving of quantitative GRE problems: A mere effort interpretation. *Personality and Social Psychology Bulletin*, *35*, 1301–1314. doi:10.1177/0146167209335165

Jamieson, J. P., & Harkins, S. G. (2011). The intervening task method: Implications for measuring mediation. *Personality and Social Psychology Bulletin*, *37*, 652–661. doi:10.1177/0146167211399776

Jamieson, J. P., Harkins, S. G., & Williams, K. D. (2010). Need threat can motivate performance after ostracism. *Personality and Social Psychology Bulletin*, *36*, 690–702. doi:10.1177/0146167209358882

Kane, M. J., Bleckley, M. K., Conway, A. R. A., & Engle, R. W. (2001). A controlled-attention view of working memory capacity. *Journal of Experimental Psychology: General*, *130*, 169–183. doi:10.1037/0096-3445.130.2.169

Klauer, K. C., Herfordt, J., & Voss, A. (2008). Social presence effects on the Stroop task: Boundary conditions and an alternative account. *Journal of Experimental Social Psychology*, *44*, 469–476. doi:10.1016/j.jesp.2007.02.009

Lang, P. J., & Bradley, M. M. (2008). Appetitive and defensive motivation is the substrate of emotion. In A. J. Elliott (Ed.), *Handbook of approach and avoidance motivation* (pp. 51–65). New York, NY: Psychology Press.

Latané, B. (1981). The psychology of social impact. *American Psychologist*, *36*, 343–356. doi:10.1037/0003-066X.36.4.343

MacKinnon, D. P., Geiselman, R. E., & Woodward, J. A. (1985). The effects of effort on Stroop interference. *Acta Psychologica*, *58*, 225–235. doi:10.1016/0001-6918(85)90022-8

Markus, H. (1978). The effect of mere presence on social facilitation: An unobtrusive test. *Journal of Experimental Social Psychology*, *14*, 389–397. doi: 10.1016/0022-1031(78)90034-3

Markus, H. (1981). The drive for integration: Some comments. *Journal of Experimental Social Psychology*, *17*, 257–261. doi:10.1016/0022-1031(81)90026-3

McFall, S. R., Jamieson, J. P., & Harkins, S. G. (2009). Testing the generalizability of the mere effort account of the evaluation-performance relationship. *Journal of Personality and Social Psychology*, *96*, 135–154. doi:10.1037/a0012878

Muller, D., & Butera, F. (2007). The focusing effect of self-evaluation threat in coaction and social comparison. *Journal of Personality and Social Psychology*, *93*, 194–211. doi: 10.1037/0022-3514.93.2.194

Normand, A., Bouquet, C. A., & Croizet, J. (2014). Does evaluative pressure make you less or more distractible? Role of top-down attentional control over response selection. *Journal of Experimental Psychology: General*, *143*, 1097–1111. doi:10.1037/a0034985

O'Brien, L. T., & Crandall, C. S. (2003). Stereotype threat and arousal: Effects on women's math performance. *Personality and Social Psychology Bulletin*, *29*, 782–789. doi: 10.1177/0146167203029006010

O'Malley, J. J., & Poplawsky, A. (1971). Noise-induced arousal and breadth of attention. *Perceptual and Motor Skills*, *33*, 887–890. doi:10.2466/pms.1971.33.3.887

Pallak, M. S., Pittman, T. S., Heller, J. F., & Munson, P. (1975). The effect of arousal on Stroop color-word task performance. *Bulletin of the Psychonomic Society*, *6*, 248–250.

Paulus, P. B., & Murdoch, P. (1971). Anticipated evaluation and audience presence in the enhancement of dominant responses. *Journal of Experimental Social Psychology*, *7*, 280–291. doi:10.1016/0022-1031(71)90028-X

Pessin, J. (1933). The comparative effects of social and mechanical stimulation on memorizing. *American Journal of Psychology*, *45*, 263–270. doi:10.2307/1414277

Platania, J., & Moran, G. P. (2001). Social facilitation as a function of mere presence of others. *Journal of Social Psychology*, *141*, 190–197. doi:10.1080/00224540109600546

Queralt, A., Weerdesteyn, V., van Duijnhoven, H., Castellote, J., Valls-Solé, J., & Duysens, J. (2008). The effects of auditory startle on obstacle avoidance during walking. *Journal of Physiology*, *18*, 4453–4463. doi:10.1113/jphysiol.2008.156042

Quinn, D. M., & Spencer, S. J. (2001). The interference of stereotype threat with women's generation of mathematical problem-solving strategies. *Journal of Social Issues*, *57*, 55–71. doi:10.1111/0022-4537.00201

Roberts, R. J., Hager, L. D., & Heron, C. (1994). Prefrontal cognitive processes: Working memory and inhibition in the antisaccade task. *Journal of Experimental Psychology: General*, *123*, 374–393. doi:10.1037/0096-3445.123.4.374

Rohleder, N., Wolf, J. M., Maldonado, E. F., & Kirschbaum, C. (2006). The psychosocial stress-induced increase in salivary alpha-amylase is independent of saliva flow rate. *Psychophysiology*, *43*, 645–652. doi:10.1111/j.1469-8986.2006.00457.x

Sanna, L. J. (1992). Self-efficacy theory: Implications for social facilitation and social loafing. *Journal of Personality and Social Psychology*, *62*, 774–786. doi: 10.1037/0022-3514.62.5.774

Schmader, T., & Johns, M. (2003). Converging evidence that stereotype threat reduces working memory capacity. *Journal of Personality and Social Psychology*, *85*, 440–452. doi: 10.1037/0022-3514.85.3.440

Schmitt, B. H., Gilovich, T., Goore, N., & Joseph, L. (1986). Mere presence and social facilitation: One more time. *Journal of Experimental Social Psychology*, *22*, 242–248. doi: 10.1016/0022-1031(86)90027-2

Seitchik, A. E., & Harkins, S. G. (2014). The effects of nonconscious and conscious goals on performance. *Basic and Applied Social Psychology*, *36*, 99–110. doi: 10.1080/01973533.2013.856785

Seitchik, A. E., & Harkins, S. G. (2015). Stereotype threat, mental arithmetic, and the mere effort account. *Journal of Experimental Social Psychology*, *61*, 19–30. doi:10.1016/j.jesp.2015.06.006

Seitchik, A. E., Jamieson, J., & Harkins, S. G. (2014). Reading between the lines: Subtle stereotype threat cues can motivate performance. *Social Influence*, *9*(1), 52–68. doi:10.1080/15534510.2012.746206

Sereno, A. B. (1992). Programming saccades: The role of attention. In K. Rayner (Ed.), *Eye movements and visual cognition* (pp. 89–107). New York, NY: Springer.

Sharma, D., Booth, R., Brown, R., & Huguet, P. (2010). Exploring the temporal dynamics of social facilitation in the Stroop task. *Psychonomic Bulletin & Review*, *17*, 52–58. doi:10.3758/PBR.17.1.52

Spence, K. W. (1956). *Behavior theory and conditioning*. New Haven, CT: Yale University Press. doi:10.1037/10029-000

Stroebe, W. (2012). The truth about Triplett (1898), but nobody seems to care. *Perspectives on Psychological Science*, *7*, 54–57. doi:10.1177/1745691611427306

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, *18*, 643–662. doi:10.1037/h0054651

Stuyven, E., Van der Goten, K., Vandierendonck, A., Claeys, K., & Crevits, L. (2000). The effect of cognitive load on saccadic eye movements. *Acta Psychologica*, *104*, 69–85. doi:10.1016/S0001-6918(99)00054-2

Szymanski, K., Garczynski, J., & Harkins, S. G. (2000). The contribution of the potential for evaluation to coaction effects. *Group Processes & Intergroup Relations*, *3*, 269–283. doi:10.1177/1368430200033003

Szymanski, K., & Harkins, S. (1987). Social loafing and self-evaluation with a social standard. *Journal of Personality and Social Psychology*, *53*, 891–897. doi:10.1037/0022-3514.53.5.891

Szymanski, K., & Harkins, S. (1993). The effect of experimenter evaluation on self-evaluation within the social loafing paradigm. *Journal of Experimental Social Psychology*, *29*, 268–286. doi:10.1006/jesp.1993.1012

Thomas, S. L., Skitka, L. J., Christen, S., & Jurgena, M. (2002). Social facilitation and impression formation. *Basic and Applied Social Psychology*, *24*, 67–70. doi:10.1207/153248302753439146

Travis, L. (1925). The effect of a small audience upon eye-hand coordination. *Journal of Abnormal and Social Psychology*, *20*, 142–146. doi:10.1037/h0071311

Travis, L. (1928). The influence of the group upon the stutterer's speed in free association. *Journal of Abnormal and Social Psychology*, *23*, 45–51. doi:10.1037/h0074512

Triplett, N. (1898). The dynamogenic factors in pacemaking and competition. *American Journal of Psychology*, *9*, 507–533. doi:10.2307/1412188

Unsworth, N., Schrock, J. C., & Engle, R. W. (2004). Working memory capacity and the antisaccade task: Individual differences in voluntary saccade control. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *30*, 1302–1321. doi: 10.1037/0278-7393.30.6.1302

Utman, C. H., & Harkins, S. G. (2010). The effect of increasing ego-involvement on the potency of the potential for self-evaluation. *Journal of Applied Social Psychology*, *40*, 1579–1604. doi:10.1111/j.1559-1816.2010.00630.x

White, P. H., Kjelgaard, M. M., & Harkins, S. G. (1995). Testing the contribution of self-evaluation to goal-setting effects. *Journal of Personality and Social Psychology*, *69*, 69–79. doi:10.1037/0022-3514.69.1.69

Williams, K. D. (2009). Ostracism: A temporal need-threat model. In M. Zanna (Ed.), *Advances in experimental social psychology* (pp. 279–314). New York, NY: Academic Press. doi:10.1016/S0065-2601(08)00406-1

Witte, K. L., & Freund, J. S. (2001). Single-letter retrieval cues for anagram solution. *Journal of General Psychology*, *128*, 315–328. doi:10.1080/00221300109598914

Wuhr, P., & Huestegge, L. (2010). The impact of social presence on voluntary and involuntary control of spatial attention. *Social Cognition*, *28*, 145–160. doi:10.1521/soco. 2010.28.2.145

Yantis, S., & Jonides, J. (1990). Abrupt visual onsets and selective attention: Voluntary versus automatic allocation. *Journal of Experimental Psychology: Human Perception and Performance*, *16*, 121–134. doi:10.1037/0096-1523.16.1.121

Zajonc, R. B. (1965). Social facilitation. *Science*, *149*, 269–274. doi:10.1126/science. 149.3681.269

Zajonc, R. B. (1980). Compresence. In P. Paulus (Ed.), *Psychology of group influence* (pp. 35–60). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Allison E. Seitchik

Department of Psychology, Merrimack College

Adam J. Brown

Department of Psychology, Northeastern University

Stephen G. Harkins

Department of Psychology, Northeastern University