

The effect of spectator presence on gymnastic performance in a field situation*

PAUL B. PAULUS, JOHN C. SHANNON, DEXTER L. WILSON
and THOMAS D. BOONE
University of Texas at Arlington, Arlington, Tex. 76010

The effect of spectator presence on gymnastic performance was examined in a series of studies. It was found that spectator presence led to a significant decrement in quality of gymnastic performance when Ss were given a forewarning about the presence of the spectators but had no overall effect when Ss were not forewarned. This finding provides support for the anticipated evaluation modification of the Zajonc (1965) hypothesis of audience effects (cf. Paulus & Murdoch, 1971). Furthermore, the relationship between the initial level of skill of the Ss and the change in their performance during spectator presence in both studies suggested the need for additional modification of the Zajonc position.

Recent laboratory research has shown a renewed interest in the effects of spectator presence on task performance (cf. Cottrell, 1972). The present study was concerned with the extent to which predictions derived from this research can be extended to behavior in field situations. One such situation, in which spectator presence is an important factor, is athletics. While athletic practice usually takes place in the presence of few, if any, spectators, actual athletic contests often involve large numbers of spectators. Although a "naive" analysis of the situation would suggest that athletes should "try harder" in front of spectators and hence perform better under such conditions than in practice, recent laboratory research (e.g., Cottrell, Rittle, & Wack, 1967; Martens, 1969) contradicts such a simplistic position. These studies support a prediction by Zajonc (1965) that the presence of spectators is a source of drive arousal and consequently facilitates performance on well-learned tasks but interferes with performance on poorly learned tasks. This prediction is based on the Hull-Spence theory proposal that drive arousal enhances the emission of dominant responses relative to subordinate ones (cf. Spence, 1956). On well-learned tasks, correct responses are supposedly dominant, while the incorrect response tendencies are presumably dominant on poorly learned tasks.

In the present series of studies, the effect of spectator presence on a gymnastic routine was assessed. The individual parts of the routine had

been practiced for about 8 weeks, while the routine as a unit had been practiced on the 2 class days prior to the experimental sessions. While one might expect the correct responses to be dominant for most of these students, no independent evidence of this was available. An unambiguous prediction about the overall effect of spectator presence on performance in such a situation is not possible. However, if one assumes greater dominance of the correct response tendencies for the relatively skilled Ss than for the relatively unskilled ones, one can predict that the skilled Ss should exhibit less of a decrement in performance or more of an increment of performance in the presence of spectators than would unskilled Ss, depending on whether the correct responses were subordinate or dominant, respectively.

EXPERIMENT 1

Procedure

Twenty-five male students from an introductory gymnastic class at the University of Texas at Arlington served as Ss in the experiment. On the first class day, all Ss were asked to perform a floor exercise routine which consisted of a combination of forward and backward rolls, a Swedish fall, three leg circles, a headstand, and a dive roll. The Ss were instructed to go individually to the gymnastic court in an alphabetical order and to perform their routine to the best of their ability. They were told that this type of individualized practice had been helpful in the past semester and was being tried again this semester for that reason. The instructor indicated that he would explain the reasons in more detail later. No spectators were allowed in the area where the routine was performed. The only individual present was an experimental assistant who was hidden behind a set of parallel bars which were covered with floormats in a dimly lit corner of the

court. The assistant recorded the performance of the Ss with a video tape recorder which was not visible to the Ss.

On the second class day, 12 of the Ss were again asked to perform their routines alone (alone-alone condition). The other 13 were asked to perform their routines in front of an audience of 17 male and female student volunteers, who sat on the floor next to the exercise mat used for the routine (alone-audience condition). They sat quietly and did not respond overtly in any way to the gymnastic performance. The alone-audience Ss were not informed of the spectators' presence until they entered the gymnastic court for their individual routines. At this time, each S was told privately of the spectator presence and asked to perform his routine as best he could. The S then performed his routine immediately. Each S was sent to the locker room after his performance. All of the Ss in the alone-alone condition were run prior to those in the alone-audience condition.

The Ss in the alone-alone and alone-audience conditions were matched on the basis of their performance on the first test day. Three gymnastic instructors rated the goodness of their performance from the video tape on a 9-point scale. All Ss were paired with someone of a similar rating, and one S from each pair was assigned randomly to either of the two experimental conditions.

The dependent variable was the rating of the individual performance provided by two experienced judges. One of them was qualified as a judge for gymnastic meets by the U.S. Gymnastic Federation, while the other was trained in judging by the first judge. They deducted a standard number of points from a possible total of 15 points for flaws in the execution of the routine. The exact number of points deducted for each type of flaw was based on standards provided by the International Gymnastic Federation.¹

Results

The correlation between the ratings of the two judges was .79. This correlation is statistically significant ($p < .001$). The average rating of the two judges for each S was used as the dependent variable. The mean ratings for each condition are shown in Table 1. These ratings indicate that the Ss were relatively proficient at the task in that they received about two-thirds of the possible 15 points. On this basis, one might assume that their correct responses were dominant in the task situation and might predict that audience presence should enhance their performance. Contrary to these

*This research was supported by funds provided by the Liberal Arts Organized Research Committee at the University of Texas at Arlington. The help of Bill Cornelius, Rick Church, Nancy Mayes, Allison Sloan, and Gary Tonn in judging the gymnastic routine is gratefully acknowledged.

Table 1
Mean Performance Ratings for Experiment I

	Condition	
	Alone- Alone	Alone- Audience
Day 1	9.86	9.59
Day 2	10.62	10.13

expectations, the only significant effect of the two-factor repeated-measures analysis of variance was a main effect for days ($F = 6.94$, $df = 1,23$, $p < .025$). This reflects the fact that Ss in both the alone-alone and the alone-audience conditions improved about equally in performance on the second day. Further disconfirmation of the predictions derived from Zajonc's theory was provided by a highly negative correlation ($r = -.83$, $df = 10$, $p < .001$) between the performance scores of the alone-audience Ss on the first day and the amount of improvement they exhibited on the second day.² Thus, the higher the level of initial skill, the smaller the amount of improvement. A similar but nonsignificant trend occurred in the alone-alone condition ($r = -.38$), indicating that a ceiling effect might account for part of the correlation in the alone-audience condition.

EXPERIMENT 2

It is possible that one feature of the procedure in the alone-audience condition of Experiment 1 may have been an important determinant of the failure to obtain an overall audience effect. All Ss in this condition were informed of the spectator presence only when their turn came to perform. Results of several studies (e.g., Breznitz, 1967; Folkins, 1970; Nomikos, Opton, Averill, & Lazarus, 1968) have suggested that only minimal levels of arousal are produced by a fear-arousing stimulus when the individual is given no forewarning of the impending presentation of this stimulus. A short warning period appears to lead to greater degrees of arousal. That this time variable may be important with audience presence as well is suggested by recent theoretical modifications of Zajonc's (1965) proposal (cf. Cottrell, 1972; Henchy & Glass, 1968; Paulus & Murdoch, 1971). These propose that anticipation of evaluation and not mere spectator presence is responsible for audience effects on performance. The evaluation interpretation suggests that the individual's anticipation of his performance is the important mediator of drive arousal. One could argue that presenting an audience without prior warning would limit the occurrence of an anticipatory process. Consequently,

allowing the S some period of time for the anticipation to occur may be necessary to demonstrate the audience effects. This would be especially important in situations where the performance is very brief, as in the present study (about 30 sec). Situations involving somewhat longer performance trials would allow sufficient time for the elicitation of evaluation anticipations during the actual performance period. The second experiment provides a test of the hypothesis that some degree of anticipation time may be necessary for audience effects to occur.

Method

This experiment was an exact replication of Experiment 1 with two exceptions—the audience manipulation and the S population. On the second day in the alone-audience condition, all of the Ss were told as a group of the presence of the spectators after the alone-alone Ss had finished their performance. About 1 min later, the first S began his performance in front of 17 spectators. Thus, all of the alone-audience Ss were given some degree of forewarning, the exact amount being dependent on their places in the random order of running (ranging from 1 to 10 min). The Ss employed in this study were 21 male and female students in an advanced gymnastic class. Ten performed in the alone-audience condition, while 11 performed in the alone-alone condition. Since this group was assumed to be superior to that of Experiment 1, it was expected that their correct responses would be dominant in the task situation. It was predicted that spectator presence would enhance the gymnastic performance and would produce a positive correlation between task improvement and level of initial skill.

Results

Again, the dependent measure was the average rating of the two judges. The correlation between the two sets of ratings was .60. This is statistically reliable ($p < .01$). The means for the ratings are shown in Table 2. The Day 1 performance ratings of these Ss were significantly superior to those of Experiment 1 ($t = 3.09$, $df = 44$, $p < .01$). There was a tendency for the Ss to perform more poorly on the second day than on the first. This is reflected in a significant main effect of days ($F = 24.43$, $df = 1,19$, $p < .001$). This performance decrement was greater for the alone-audience Ss than for the alone-alone Ss, as indicated by a significant Audience Condition by Day interaction ($F = 7.38$, $df = 1,19$, $p < .025$). Specific comparisons of the Day 1 and Day 2 scores for both the alone-alone and alone-audience conditions indicated that only the

decrement in the alone-audience condition was statistically significant (alone-audience, $F = 11.62$, $df = 1,19$, $p < .005$; alone-alone, $F = 1.13$, $df = 1,19$, n.s.). The within-cell correlation of the initial performance scores with the amount of decrement in performance on the second day for the alone-audience group was highly positive ($r = .84$, $df = 8$, $p < .01$). That is, the higher the level of skill of the S, the greater the decrement in performance produced by audience presence. The correlation in the alone-alone condition was smaller, but approached significance ($r = .55$, $df = 9$, $p < .10$). Thus, the correlation in the alone-audience condition may be partly due to a regression effect.

DISCUSSION

The performance decrement in the alone-audience condition and the finding that this decrement increased with higher levels of skill are both counter to Zajonc's Hull-Spence analysis of audience effects. These findings might be explained by utilizing some of the implications of the Broen & Storms (1961) extension of Hull-Spence theory.³ They propose that response tendencies may have a response-strength ceiling. Drive arousal should tend to raise the strength of the dominant response toward its ceiling, at the expense of the subordinate responses. With very high levels of drive arousal, however, the dominant response should reach its ceiling and further increases should occur only in the strength of the competing subordinate responses. As a result, the dominance of the dominant response relative to the subordinate responses decreases and the probability of occurrence of the subordinate responses increases. This theory predicts an inverted-U relationship between level of drive arousal and the performance of a complex task in which the correct response is dominant. Thus, in the present study, providing an individual a forewarning of the spectator presence may have produced a relatively high level of drive arousal and consequently a decrement in performance. Spectator presence without a forewarning, as in Experiment 1, may have produced only a moderate level of drive arousal, and hence no decrement in performance occurred.⁴

Table 2
Mean Performance Ratings for Experiment II

	Condition	
	Alone- Alone	Alone- Audience
Day 1	10.77	11.42
Day 2	10.31	9.89

This type of analysis can also account for the correlations between initial level of skill and the change in performance during audience presence in Experiments 1 and 2. In Experiment 1, it was found that relatively unskilled Ss in the alone-audience condition showed more improvement in performance than relatively skilled Ss. One can assume that the dominant correct responses with respect to the various parts of the routine were closer to their response strength ceiling for the relatively skilled Ss than for the less skilled Ss. Although a moderate level of drive arousal may have facilitated the dominant correct tendencies for all Ss, only those of the skilled Ss may have reached their response strength ceiling. However, the continued facilitation of the competing incorrect response tendencies, once this ceiling had been reached, may have led to less overall improvement for these Ss.

In Experiment 2, a greater decrement in performance was exhibited by relatively skilled Ss than relatively unskilled Ss. Since this study involved somewhat higher levels of overall skill and a supposedly higher level of audience-produced arousal, the dominant correct tendencies may have been brought to their response strength ceiling for all Ss. However, the continued facilitation of the competing incorrect tendencies, once this ceiling was reached, may have led to a resultant decrement in performance. This decrement may have been less for the relatively unskilled than for the relatively skilled Ss, since more facilitation of the correct dominant response was possible for the unskilled Ss.

It could be argued that, since the Ss in Experiment 2 were more skilled than those in Experiment 1, the Broen and Storms analysis could account for the data without assuming differential drive level in the two studies. Two factors suggest that this is not the case. First of all, the performance ratings employed in these experiments actually involve the summation of the ratings of two distinct aspects of performance, style, and the degree to which the individual performs all of the required parts of the routine. While the two classes differed significantly in terms of the style rating, they did not differ significantly

in execution rating. An analysis of the effect of audience presence on the execution ratings entirely parallels the analysis reported in this paper. Hence, difference in overall skill level cannot by itself explain the results.

A third experiment, done in a prior semester with introductory students, further reinforces this conclusion. In that study, 18 students from an introductory gymnastic class performed their routines alone on one day and in front of an audience the next day. The audience consisted of about 20 volunteers and the entire gymnastic class. Thus, an anticipation period was provided for all Ss in this study. Two judges rated the quality of performance on a 10-point scale in both sessions. (They were unobtrusive but present in the same room on the first day.) It was found that audience presence led to a significant decrement in performance ($F = 4.48$, $df = 1,16$, $p \approx .05$). In addition, this decrement was greater with the more highly skilled students, as indicated by a positive correlation between initial level of skill and amount of decrement in performance on the second day ($r = .80$, $df = 16$, $p < .001$). The results of this study are entirely in accord with those of Experiment 2, despite the fact that this experiment did not have some of the control features of the previous experiments.

Several implications of the above studies should be pointed out. First, they suggest that in order to predict the quality of performance on a task in front of spectators one must determine both the degree to which incorrect competing response tendencies are elicited during task performance (level of skill) and the level of drive arousal that is produced by a particular spectator situation (cf. Broen & Storms, 1961). Secondly, the present group of studies is the first in the author's knowledge that has used a "pure" alone condition. In all previous audience studies, the E was always present or monitoring the performance nearby. In the present study, no one was visibly present, and there was no obvious indication that performance was being monitored. Finally, this series of studies has used a relatively large number of spectators. Most previous studies have used only one or two. Since the Broen & Storms (1961) analysis suggests an inverted-U

relationship between quality of performance on a complex task and level of drive arousal, a study assessing the effect of the number of spectators on the performance of such a task would be quite informative.

REFERENCES

- BREZNITZ, S. Incubation of threat: Duration of anticipation and false alarm as determinants of the fear reaction to an unavoidable frightening event. *Journal of Experimental Research in Personality*, 1967, 2, 173-179.
- BROEN, W. E., JR., & STORMS, L. H. A reaction potential ceiling and response decrements in complex situations. *Psychological Review*, 1961, 68, 405-415.
- COTTRELL, N. B. Social facilitation. In C. G. McClintock (Ed.), *Experimental social psychology*. New York: Holt, Rinehart & Winston, 1972.
- COTTRELL, N. B., RITTLE, R. H., & WACK, D. L. The presence of an audience and list type (competitional or noncompetitional) as joint determiners of performance in paired associate learning. *Journal of Personality*, 1967, 35, 425-434.
- FOLKINS, C. H. Temporal factors and the cognitive mediators of stress reaction. *Journal of Personality & Social Psychology*, 1970, 14, 173-184.
- HENCHY, T., & GLASS, D. C. Evaluation apprehension and the social facilitation of dominant and subordinate responses. *Journal of Personality & Social Psychology*, 1968, 10, 446-454.
- MARTENS, R. Effect of an audience on learning and performance of a complex motor skill. *Journal of Personality & Social Psychology*, 1969, 12, 252-260.
- NOMIKOS, M. S., OPTON, E. M., JR., AVERILL, J. R., & LAZARUS, R. S. Surprise versus suspense in the production of stress reaction. *Journal of Personality & Social Psychology*, 1968, 2, 204-208.
- PAULUS, P. B., & MURDOCH, P. Anticipated evaluation and audience presence in the enhancement of dominant responses. *Journal of Experimental Social Psychology*, 1971, 7, 280-291.
- SPENCE, K. W. *Behavior theory and conditioning*. New Haven, Conn: Yale University Press, 1956.
- ZAJONC, R. B. Social facilitation. *Science*, 1965, 149, 269-274.

NOTES

1. The presence of spectators was not evident from video tapes. Consequently, the judges did not know which Ss had performed in front of spectators.
2. A constant was added to the change scores in this correlational analysis and in the subsequent ones reported in order to make all change scores positive.
3. This interpretation was suggested to us by Verne C. Cox.
4. An analysis of the relationship between amount of anticipation time and amount of response decrement in the alone-audience condition of Experiment 2 by Spearman rho indicated this relationship was minimal ($\rho = -.06$). This suggests that the important factor is no warning vs some warning, not the amount of warning time.