

A FACTORIAL STUDY OF PRAISE AND BLAME¹

WALLACE A. KENNEDY AND HERMAN C. WILLCUTT

Florida State University

Summary.—The effectiveness of praise and blame on the discrimination performance of 720 Ss as a function of race (Negro, white), grade (2, 4, 7, 10), intelligence (high, medium, low), and sex was studied. Analysis of variance indicated that in situations involving perceptual skills, with speed as the criterion measure, reproof had a debilitating effect upon performance, while praise resulted in an increase in performance larger than that associated with practice alone.

Several broad reviews of the extensive philosophical and empirical literature on the relative effectiveness of praise and blame are available (Hurlock, 1931; Davis & Ballard, 1932; Brenner, 1934; Murphy, Murphy, & Newcomb, 1937, pp. 376-508; Schmidt, 1941; Kennedy & Willcutt, 1964). These reviews have emphasized a lack of uniformity in methodology and conditions among these studies and the resulting muddled picture of the relative effectiveness of these incentives. The present study, employing 720 Ss, the variables of grade, intelligence, sex, and race, as well as variability between examiners, socio-economic levels, and schools, was undertaken because the effect of praise and blame appeared extremely variable and highly reliant on subject and situational modifications. Use of a larger sample in systematic examination of variables described in the literature should make possible a clearer conception of the effectiveness of praise and blame in the classroom.

METHOD

Subjects

The 720 Ss, equally divided as to race, sex, grade, and intelligence into three verbal reward condition groups, were drawn from a large southern metropolitan area (population 314,000). Using the 1960 census as a guide line, the parents of this sample were characterized as being in higher occupational levels and as having a higher level of education than the general population, a fact which should be kept in mind when generalizing from these data.

Materials

The total IQ score of the Short Form A of the California Mental Maturity Test, because of its wide use in the school systems of the area and because it provided the gross classification needed, was used to divide Ss of each racial group into three levels for intelligence. The decision to divide the sample by race be-

¹The research reported herein was supported through the Cooperative Research Program of the Office of Education, U. S. Department of Health, Education, and Welfare (C.R.P. 1929).

fore dividing by intelligence was based upon the 1963 study by Kennedy, *et al.*, which points out that the mean for Southeastern Negro children on nationally standardized examinations of intelligence is 80.1, with a *SD* of 11. Since it must be assumed that the intellectual parameter is distributed normally for the Negro population, the most reasonable solution to the problem of representativeness was to sample randomly from both populations and divide the groups into thirds. For the white sample the mean for the group was 97; *SD*, 17. The high group range was from 110 to 136; the medium range, 97 to 109; and the low, 60 to 96. For the Negro sample the mean for the group was 82; *SD*, 13. The high group ranged from 89 to 132; the medium group, from 78 to 88; and the low, from 60 to 77.

The four-key discrimination box for *S* and a panel through which was recorded *S*'s response latency in 100ths of a second, were identical to those used by Kennedy, Turner, and Lindner (1962) and Willcutt and Kennedy (1963). The discrimination task, also used in these two previous studies, consisted of 32 oddity-problem stimulus cards presenting four patterns, one of which was different from the other three. This has been described in detail by Turner (1962).

Procedure

Each *S* was individually administered two trials on the discrimination task, i.e., *S* saw twice all 32 stimulus cards in the same order. Between trials the experimental verbal incentive condition was presented.

At the end of Trial 1 blame was administered to *S*s in the blame group as follows:

Well, that's it. (Pause.) Your scores are not nearly as good as I thought they would be. I'm really disappointed. I think these are the worst scores I've gotten yet. (Pause.) But perhaps if we ran through the cards again. . . I'll tell you what. Just forget these scores. Let's go through the cards again. Do you understand the directions? Let me repeat them once more.

The directions were then reread verbatim and the same 32 cards readministered in the same order.

*S*s in the praise group were told:

Well, that's it. (Pause.) Your scores are very good. I knew you would do well, but you did even better than I expected. I am really pleased. These look like the best scores I have gotten yet. (Pause.) Now, I'd like you to go through the cards again. To get the time the same, I'll repeat the directions.

*S*s in the control group were asked questions of a biographical nature during the interval between trials. The time to do this was equated as closely as possible with that used to present the verbal incentives.

Criterion Measure

The task could be solved by all *S*s, however, the speed of solution was highly variable. These latency data, the reaction time between the appearance of the

stimulus card on S's viewing screen and his depression of the response key, were subjected to analysis.

Design

Lindquist's repeated measures analysis of variance was used to analyze data in terms of reward condition, grade level, level of intelligence, and sex. Because of the difficulty in equating for intelligence, Negro and white data were analyzed separately. There were three reward conditions: praise, blame, control; four grade levels: 2, 4, 7, 10; and three levels of intelligence: high, medium, low. There were 5 Ss in each of 144 cells. Also, in an attempt to assess the stability of the measure, the data were analyzed as a function of socio-economic level, examiner variability, and variability among participating schools.

RESULTS

Two separate analyses were run due to the number of variables contributing to the between- and within-subject error estimates and the limited number of Ss in individual cells. The first included reward, trial, grade, and intelligence; the second, reward, trial, grade, and sex. Since the major interest of this study was the effect of the verbal incentive on Ss' performance, incidental findings are not reported. Attention is focused solely on praise and blame effectiveness on performance.

White Sample

As illustrated in Table 1, the main effect of verbal incentive was nonsignificant for the white sample. However, there was significant interaction between reward and trial, between reward, trial, and grade, and between reward, trial,

TABLE 1
FOR WHITE SAMPLE: ANALYSIS OF VARIANCE AMONG FACTORS OF REWARD,
TRIAL, GRADE, AND INTELLIGENCE

Source	<i>df</i>	<i>MS</i>	<i>F</i>	<i>p</i>
Reward (R)	2	1,100.93	.932	
Trial (T)				
Grade (G)				
Intelligence (I)				
R by T	2	6,080.23	36.611	<.001
R by G	6	1,214.54	1.028	
R by I	4	1,891.63	1.601	
R by T by G	6	367.64	2.214	<.05
R by T by I	4	195.70	1.178	
R by G by I	12	1,212.39	1.026	
R by T by G by I	12	416.23	2.506	<.001
Error between	324	1,181.24		
Error within	324	166.08		
Total	719			

TABLE 2
SIGNIFICANT MEAN REACTION TIMES FOR REWARD, TRIAL, GRADE, AND
INTELLIGENCE FOR WHITE SAMPLE

G by I	Praise Trials			Blame Trials			Control Trials		
	1	2	Diff.	1	2	Diff.	1	2	Diff.
Grade 2									
High	6.97	6.48	-.49	7.12	7.74	.62	7.43	6.61	-.82
Medium	10.10	7.85	-2.25	7.38	7.73	.35	7.58	7.00	-.58
Low	8.44	7.41	-1.03	8.01	7.15	-.86	8.63	8.00	-.63
Total	8.51	7.25	-1.26	7.50	7.54	.04	7.88	7.20	-.68
Grade 4									
High	4.99	3.66	-1.33	3.87	5.37	1.50	4.43	4.00	-.43
Medium	6.03	4.85	-1.18	5.23	5.23	0.00	4.58	4.02	-.56
Low	7.87	6.05	-1.82	5.75	5.91	.16	7.26	6.25	-1.01
Total	6.30	4.85	-1.45	4.95	5.50	.55	5.42	4.76	-.61
Grade 7									
High	4.23	3.26	-.97	3.03	2.50	-.53	3.69	2.83	-.86
Medium	3.93	3.23	-.70	4.96	7.12	2.16	4.14	3.62	-.52
Low	5.09	3.86	-1.23	5.41	8.66	3.25	8.21	5.88	-2.33
Total	4.42	3.45	-.97	4.47	6.10	1.63	5.35	4.11	-1.24
Grade 10									
High	4.45	3.67	-.78	3.51	3.69	.18	4.13	3.27	-.86
Medium	5.29	4.66	-.63	5.25	5.62	.37	4.77	3.58	-1.19
Low	5.34	4.14	-1.20	6.37	7.65	1.28	6.11	4.91	-1.20
Total	5.03	4.15	-.88	5.04	5.65	.61	5.00	3.92	-1.08
Total	6.06	4.93	-1.13	5.49	6.19	.70	5.91	5.00	-.91

grade, and intelligence. Table 2 illustrates the significant mean reaction times for these white Ss.

For this sample the interaction between verbal incentive and performance on Trials 1 and 2 was significant at the .01 level. Under praise and control the white children showed a decreased mean reaction time from Trial 1 to Trial 2 (-1.13 and $-.91$, respectively). Under blame Trial 2 mean reaction time increased (.70).

The significance of the interaction between rewards, trials, and grades seems to be due mainly to the fact that praise appeared to have its greatest incremental effect for the second and fourth grade students whereas blame had its greatest decremental effect for seventh and tenth grade students. Also, seventh and tenth grade students seemed to improve their scores much more under control conditions than did the second and fourth grade students in spite of the fact that their over-all reaction time was much shorter.

In examining the analysis for location of the reward by trial by grade by intelligence interaction, one can easily see that the interaction described above is sharpened when one makes a comparison between the medium intelligent sec-

ond and fourth grade Ss who made great improvements in their scores under praise. The seventh and tenth grade subjects of medium intelligence showed little gain. At the same time it was Ss of low intelligence in the seventh and tenth grades who provided the major decrease in reaction time under the control condition. It can also be noted that in the upper grades the largest difference occurred between intelligence groups in their reaction to blame. Thus, while there was no consistent pattern of differences among the three intellectual groups in the second and fourth grades, at the seventh and tenth grades the lower and medium intellectual groups were most adversely affected by blame.

Negro Sample

Table 3 illustrates that for the Negro sample, as for the white sample, verbal incentive is nonsignificant. However, there was significant interaction between

TABLE 3
FOR NEGRO SAMPLE: ANALYSIS OF VARIANCE AMONG FACTORS OF REWARD,
TRIAL, GRADE, AND INTELLIGENCE

Source	<i>df</i>	<i>MS</i>	<i>F</i>	<i>p</i>
Reward (R)	2	2,642.34	1.012	
Trial (T)				
Grade (G)				
Intelligence (I)				
R by T	2	8,356.26	17.188	<.001
R by G	6	3,397.34	1.302	
R by I	4	1,025.29	.393	
R by T by G	6	1,355.24	2.788	<.05
R by T by I	4	462.18	.951	
R by G by I	12	2,462.55	.944	
R by T by G by I	12	161.50	.332	
Error between	324	2,609.82		
Error within	324	486.17		
Total	719			

reward and trial, and between reward, trial, and grade. Significant mean reaction times are displayed in Table 4.

For the Negro Ss, significant results paralleled those of the white sample, with one exception: intelligence did not significantly affect Ss' reaction to the verbal incentives. The Negro Ss' reaction time on Trial 2 interacted significantly with verbal incentives, and there was a significant interaction between grade, trial, and incentive.

Under the praise and the control conditions Negro Ss' mean reaction time on Trial 2 decreased (-1.43 and -1.06 , respectively). Under blame these scores increased ($+.77$).

The significant reward by trial by grade interaction seemed to result from

TABLE 4
SIGNIFICANT MEAN REACTION TIMES FOR REWARD, TRIAL, AND GRADE,
FOR NEGRO SAMPLE

Grade	Praise Trials			Blame Trials			Control Trials		
	1	2	Diff.	1	2	Diff.	1	2	Diff.
2	7.79	7.26	-.53	7.42	6.76	-.66	7.38	6.25	-1.13
4	8.93	7.46	-1.47	7.02	7.95	.93	8.01	8.14	.13
7	6.89	5.31	-1.58	7.50	8.96	1.46	8.03	6.40	-1.63
10	8.50	6.37	-2.13	8.05	9.40	1.35	7.91	6.28	-1.63
Total	8.03	6.60	-1.43	7.50	8.27	.77	7.83	6.77	-1.06

the steady decrease in reaction time from Grade 2 to Grade 10 under praise condition and the increase in reaction time from Grade 4 to Grade 10 under blame. This latter difference is more striking since at the second grade level the blame group actually showed a decrease in reaction time from Trial 1 to Trial 2 which was an exception to the general trend in the data.

Total Sample

The over-all effect of the verbal incentives was consistent for Negro and white Ss: praise and control were beneficial to the students' performance while blame was detrimental. An interesting result was the relative effectiveness of these verbal incentives in relationship to grade. For instance, at the seventh-grade level for white and Negro Ss, blame was quite detrimental, whereas for second-graders the effect was negligible.

Analyses of variance run separately on white and Negro data yielded no significant interaction between sex and verbal incentives. Analyses of variance between schools, between examiners, and between socio-economic levels gave nonsignificant results for both whites and Negroes. Thus, none of these variables significantly affected Ss' latencies.

DISCUSSION

The purpose of this study was to shed some light on the problem of inconsistency in the relative effectiveness of praise and reproof. The source of this inconsistency has been difficult to determine, but it may well result from lack of comparability of methodology. The present study investigated the independent variables (grade, intellectual level, sex, and race) which seem to have been studied most often and appear most relevant to the practical use of the verbal incentives, praise and reproof. A repeated measures complex analysis of variance design was employed and improvement in discrimination speed was used as a measure of learning.

In general, the results of the present study point to a consistent superiority of the verbal incentive, praise, over reproof in simple discrimination learning. There seems little doubt that the major difference between the two incentives

in this study results from the increasing detrimental effect of blame as one moves up the grade levels. This detrimental effect also seems to a great extent to be more evident in children in the lower half of the group in intelligence. Thus, it appears that in the upper grades and with white children in the lower intellectual group, blame has very little to offer as an incentive perhaps because the children in this group have experienced so much of it. This relationship between effectiveness of incentive and intellectual ability does not seem to hold for Negro students.

REFERENCES

- BRENNER, B. Effect of immediate and delayed praise and blame upon learning and recall. *Teachers Coll. Contr. to Educ.*, 1934, No. 620.
- DAVIS, R. A., & BALLARD, C. R. Effectiveness of various types of classroom incentives. *Educ. Meth.*, 1932, 12, 134-145.
- HURLOCK, E. B. The psychology of incentives. *J. soc. Psychol.*, 1931, 2, 261-290.
- KENNEDY, W. A., TURNER, A. J., & LINDNER, R. Effectiveness of praise and blame as a function of intelligence. *Percept. mot. Skills.*, 1962, 15, 143-149.
- KENNEDY, W. A., VAN DE RIET, V., & WHITE, J. C., JR. A normative sample of intelligence and achievement of Negro elementary school children in the Southeastern United States. *Child Developm. Monogr.*, 1963, 28, No. 6. (Serial No. 90)
- KENNEDY, W. A., & WILLCUTT, H. C. Praise and blame as incentives. *Psychol. Bull.*, 1964, 62, 323-332.
- MURPHY, G., MURPHY, L. B., & NEWCOMB, T. M. *Experimental social psychology*. New York: Harper, 1937.
- SCHMIDT, H. O. The effects of praise and blame as incentives to learning. *Psychol. Monogr.*, 1941, 53, No. 3 (Whole No. 240).
- TURNER, A. J. Discrimination reaction time as a function of anxiety and task difficulty. Unpublished Ph.D. dissertation, Florida State Univer., 1962. (*Dissert. Abstr.*, 1962, 23, 712, No. 62-3520)
- WILLCUTT, H. C., & KENNEDY, W. A. Relation of intelligence to effectiveness of praise and reproof as reinforcers for fourth-graders. *Percept. mot. Skills*, 1963, 17, 695-697.

Accepted August 20, 1965.