DOCUMENT RESUME

ED 104 954 TM 004 412

AUTHOR Brady, Peter J.; And Others
TITLE Locus of Evaluation in Children's Learning from Textbook Material.
PUB DATE [Apr 75]
NOTE 4p.; Paper presented at the Annual Meeting of the American Educational Research Association (Washington, D.C. March 30-April 3, 1975)

EDRS PRICE MF-$0.76 HC-$1.58 PLUS POSTAGE
DESCRIPTORS Academic Achievement; Anxiety; Elementary Education; *Evaluation Methods; Grade 4; *Instructional Materials; *Learning Characteristics; Reinforcement; Response Mode; Responsibility; *Self Concept; *Self Evaluation; Sex Differences; Student Attitudes; Student Evaluation; Student Teacher Relationship; Teaching Methods

ABSTRACT

Seventy-eight fourth grade children were randomly assigned to one of two evaluation groups. One group (self-evaluation) judged the correctness of their answers and reinforced themselves while the other group (other-evaluation) was judged and reinforced by some other person. Results showed that girls were significantly more responsible for failure than boys and that in the self-evaluation condition boys were significantly less anxious than girls and less likely to lie. Moreover, the results suggested that boys who evaluated themselves as compared to boys who were evaluated by others tended to experience reduced anxiety and an enhanced self-concept. (Author)

AERA: SESSION NUMBER: C-41

Locus of Evaluation in Children's Learning from Textbook Material

Peter J. Brady, John P. Rickards
Donald W. Felker
Purdue University

Peter J. Brady,
SCC-G, Purdue University
West Lafayette,
Indiana 47907

When a student answers a teacher's question the teacher typically judges the student's answer as correct or incorrect and may even reinforce him if he has correctly responded to the question posed. The student's only contribution to this interaction is his answer. It is quite possible, however, for the teacher to more directly engage the student in the evaluation process by asking him to judge the correctness of his own answer ("Do you think your answer is correct?") and to reinforce himself ("Don't you think you did well on that question?" The former-mentioned type of evaluation (other-evaluation) involves judgment and praise of the student by another person, which in this case is the teacher. The latter-mentioned type of evaluation (self-evaluation) consists of self-judgment and self-praise.

Regarding self-praise, Felker and Thomas (1971) asked children to rank order nine positive and negative statements (e.g. positive--"I'm smarter than most kids;" Negative--"I always fail.") as statements they would typically say to themselves during school work. These investigators found that those who ranked positive statements higher than the overall class mean had significantly better self-concepts. Furthermore, Felker and Stanwyck (1971) have shown that after performing school tasks, children with high self-concepts tend to make statements about themselves that are more positive than those made by students with low self-concepts. Hence, it seems that saying positive things about oneself in relation to performance in school is one behavioral manisfestation of a high self concept. Inducing this kind of behavior (self-reinforcement) in students, therefore, could be expected to increase their self-concepts relative to students experiencing reinforcement by another person, usually the teacher. Coons and McEachern (1967) and Felker (1972) point to this possibility in their respective studies wherein students who were induced to make positive statements about themselves attained a more positive attitude towards themselves, as indicated by a significant increase in the number of endorsements of self-accepting statements.

The other component of the self-evaluation strategy, self-judgment regarding the correctness of one's answers to test questions, whould represent a less threatening academic situation than judgment of correctness by another person, which is associated with the other-evaluation strategy. Therefore, we would expect less anxiety to be present in the self-evaluation condition as compared to the other-evaluation condition.

Method. The subjects for the experiment were 78 fourth grade boys and girls. Each was randomly assigned to either the self-evaluation condition or to the other-evaluation condition. The instructional materials were taken directly from regular fourth grade textbooks in English, geography, history and mathematics. Each day a passage of text in one of the above subject matter areas was given in random order to each student. Then, four questions were asked, one at a time, about the textbook material. The same text segments and questions were given to both the self-and other-evaluation groups. The treatments were administered for eleven consecutive school days, and on the twelfth and final day of the experiment, self-concept, responsibility and anxiety measures were administered to all subjects.

The presentation of the experimental materials followed the same format each day: Instructions and a page of text (approximately 100 words) were given to each student for five minutes of study. The instructions emphasized that the students would be tested on the material presented immediately after the allotted study time. After the text page was collected the first multiple choice question was administered. Two minutes later an evaluation page was presented. In the other-evaluation group, one of three different evaluations was provided depending upon the particular answer given to a multiple choice question. These evaluations ranged from "correct" through "almost correct" to "heading in the right direction." This continuum of evaluations was chosen to insure that all evaluations in the other-evaluation condition would be at least somewhat positive and so warrant subsequent reinforcement (see praise page below) for answers given. For example, if the question was, "Is a glacier a) a snow field, b) a river of ice or, c) a lake of ice?" and the

"correct" answer was chosen, then the evaluation page was, "Your answer was correct because a glacier is a mass of ice flowing down the valley." If the "almost correct" answer was chosen, the evaluation page for the other-evaluation group was, "Your answer was almost correct because the glacier started out as snow. This snow melted and froze till eventually it was a mass of ice;" If the "heading-in-the-right-direction" answer were selected the evaluation page read "Your answer is heading in the right direction because the glacier moves so slowly it looks like a lake of ice."

Regardless of the student's answer to any given question, those in the self-evaluation group received an evaluation page containing all three explanations given above but without any reference to their degree of correctness. For example, "A glacier is a mass of ice flowing down the valley. The glacier started out as snow. This snow melted and froze till eventually it was a mass of ice. The glacier moves so slowly it looks like a lake of ice." The student was asked to examine these explanations and determine for himself if his selection of the multiple choice item was correct. Two minutes were allotted for perusal of this evaluation page after which a "praise" page was handed out. For the self-evaluation group the praise page elicited self-reinforcement through a question and answer sequence, e.g. "Don't you think you did a very good job on that question? Yes, I think I did great." For the other-evaluation group, reinforcement was given in the form of a declarative statement. "You really did well on that question. Very good!" After one minute the praise pages were collected and the second question was handed out. The sequence outlined above was then repeated for each of the remaining three questions for every day of testing.

Design. Academic performance was measured using the answers given to the textbook questions as the dependent variable. For any given question, three points were allotted for a "correct" answer, two for an "almost correct" response, and one point was allowed for an answer that was "heading in the right direction." The analysis used for the resulting data was a 2 x 11 analysis of variance with the self-evaluation and the other-evaluation conditions constituting the first variable and the eleven testing days making up the repeated measures factor.

The scales used to measure self-concept were the Piers-Harris (P-H) Self-Concept Scale (Piers, and Harris, 1964) and the Self-Concept Statement (SCS) Scale (Felker and Stanwyck, 1971 Felker and Thomas, 1971). The intellectual Achievement Responsibility (IAR) Questionnaire (Crandall, Katkovsy and Crandall, 1967) was used to measure responsibility and the Children's Manifest Anxiety Scale (CMAS) to measure anxiety (Casteneda, McCandless, and Palermo, 1956). The CMAS total scale (CMAST) contains an eleven item lie scale (CMASL). The IAR total scale (IART) has two subscales, one which measures responsibility for positive events (IAR+), and the other which measures responsibility for negative events (IAR-). Each of the above measures was examined via a 2 x 2 analysis of variance with the evaluation variable (self- and other-evaluation) and the sex variable (male and female) being the two factors involved in the analysis.

Results. The effect on academic performance of the type of evaluation employed was not significant ($p > .05$) nor was there any significant effect across the eleven days of testing. And finally, no significant ($p > .05$) interaction was present in this analysis. Although there was no significant effect on any of the affective measures due to the type of evaluation strategy used, there were some significant ($p < .05$) effects due to the sex of the students. Girls tended to score higher than boys on all scales and significantly ($p < .05$) so in the case of the IAR and CMAS scales. There were, however, no significant interactions in these analyses.

With regard to specific effects observed, girls scored significantly higher than boys on both the IART scale and the IAR- scale ($F = 5.50$, $df = 1, 77$, $p < .05$); $F = 5.23$, $df = 1,77$, $p < .05$, respectively). Moreover, girls scored higher than boys on the CMAST scale ($F = 6.85$, $df = 1, 55$, $p < .01$) and on the CMASL subscale as well ($F = 4.83$, $df = 1,77$, $p < .05$). Tests for simple main effects of sex for the self- and other-evaluation conditions were conducted using CMAS as the dependent variable. Sex differences were found in the self-evaluation group with boys ($\bar{X} = 16.33$) scoring significantly lower than girls ($\bar{X} = 21.40$) on the CMAST ($F = 4.90$, $df = 1, 38$, $p < .05$); also, boys ($\bar{X} = 3.21$) scored significantly lower than girls ($\bar{X} = 4.60$) on the lie subscale (CMASL) as well ($F = 4.49$, $df = 1, 38$, $p < .05$). While the P-H and SCS scales showed no significant ($p > .05$) effects due to evaluation or sex differences, a t test revealed that the difference between the self-evaluation condition

($\bar{X}$ = 7.27) and the other-evaluation condition ($\bar{X}$ = 8.19) approached significance for boys on the SCS scale, (t = 1.54, df = 49, $p < .13$).

Discussion. It has been found that girls typically accept responsibility for successful (IAR+) academic performance (Soloman, Houlihan and Parelius, 1969; Stanwyck, 1972) but not for unsuccessful (IAR-) academic performance. The divergence from the usual findings in the present experiment with girls scoring significantly higher than boys only on the IAR- scale was probably related to the scoring continuum that was used in this experiment. Even in the other-evaluation condition, the worst one could possible do was to choose "heading-in-the-right-direction" answers to all questions. Hence, it appears that girls who do not normally accept responsibility for failure did so, having experienced this highly success-oriented situation.

The fact that boys scored significantly lower than girls on anxiety only in the self-evaluation condition suggests that boys who were placed in a nonthreatening learning situation (as in the self-evaluation condition wherein there was no outside judgment, praise or blame) tended to be less anxious than girls and less likely to lie. On the other hand, boys and girls who were evaluated by others did not differ in their degree of anxiety or their tendency to lie. The data also hint at the possibility that boys who self-evaluate were less likely to lie (t = 1.56, df = 49, $p < .12$) than boys who were evaluated by others, the situation more typically found in our schools. Moreover, boys who evaluated themselves demonstrated a somewhat greater ($p < .13$) self-concept (SCS scale) than boys who were evaluated by others.

The above findings are especially interesting in view of the fact that three important factors which could affect any comparison between self- and other-evaluation strategies remained relatively constant in this experiment. First of all, the amount of time spent on learning by each of the groups was identical; secondly students in both the self-evaluation and other-evaluation conditions performed equally well on the academic task; thirdly, the amount of reinforcement was the same for both groups. Hence, the affective differences that did occur were likely due to the locus of judgment and reinforcement, and not simply to the fact that students who evaluated themselves experienced less failure, received more reinforcement and/or took more time than those who were evaluated by others.

Educational Implications. The distinction here between self- and other-evaluation strategies is suggestive of that between democratic and authoritarian leadership styles. The authoritarian teacher being more task-oriented would be expected to favor the other-evaluation (teacher-centered) strategy, while the democratic teacher being more person-oriented would likely favor the self-evaluation (student-centered) strategy. Research in the area (Anderson, 1959; Brophy and Good, 1974) indicates that while there is no academic performance difference due to variation in leadership style, there are relatively clear-cut affective differences. More specifically, democratic classrooms are characterized by greater student cooperation, less competitiveness and greater student enjoyment than are authoritarian classrooms (Brophy and Good, 1974). A somewhat similar pattern emerged in the present experiment: While there were no cognitive performance differences related to the type of evaluation strategy employed, there were certain affective differences. Boys who judged their own answers to questions and reinforced themselves for their performance (suggestive of democratic classrooms) as opposed to being judged and reinforced by another person (more likely associated with authoritarian classrooms) were less anxious and their tendency to lie was less than girls in the same situation. Moreover, there was some evidence that the student-centered as compared to the teacher-centered evaluation strategy reduced the boys' tendency to lie and enhanced their self-concepts. It seems then, that the affective outcomes associated with the student-centered evaluation strategy are quite harmonious with those associated with democratic classrooms. More research employing rating scales, questionnaires, observational forms in combination with various personality measures such as those used here might further reveal differences between student-centered and teacher-centered evaluation strategies.