

The experimenter bias effect: A failure to replicate

THEODORE JACOB, *University of Nebraska, Lincoln, Nebr.*
68508

The present study attempted to replicate the E-bias effect using Rosenthal's (1966) basic paradigm, examine the relationship between E-bias and Ss' perception of Es' expectancies, and evaluate the E-bias effect in relation to a no-expectancy control group. Results failed to demonstrate the E-bias effect, although Ss' judgments of Es' expectancies were related to the three different treatment groups. The present findings are discussed briefly in terms of Barber & Silver's (in press) recent review of this area.

During the past decade, a number of investigators have been examining the experimenter-subject relationship in terms of the biasing effects that experimenter characteristics exert upon forthcoming results (Kintz et al, 1965; Masling, 1966; McGuigan, 1963; Rosenthal, 1966). In particular, Rosenthal and his colleagues have gathered a considerable amount of data which ostensibly demonstrate that Es' expectancies exert a potent effect upon the results they obtain. In a recent review of the relevant literature, however, Barber & Silver (in press) re-analyzed 31 studies which attempted to demonstrate the E-bias effect and concluded that "the majority of studies did not clearly demonstrate the effect [and that] the Experimenter Bias Effect appears to be more difficult to demonstrate and less pervasive than was implied in previous reviews in this journal."

The purpose of the present study is threefold: (1) In light of the Barber and Silver paper, it appears necessary to re-evaluate the reliability of the E-bias effect by means of independent replications. (2) Although many studies by Rosenthal have incorporated rating scales and check lists to be answered by Ss after the experiment proper, most of these items have dealt with S's perception of E's personality traits and characteristics but have not probed directly S's perception of E's expectancies or desires for certain results (Rosenthal, 1966). The present study included an S questionnaire which was designed to evaluate this question. (3) With rare exception, studies of E-bias have failed to include a no-expectancy control group (Barber et al, 1966; Friedman, 1964). If the E-bias effect is obtained, omission of a no-expectancy control would leave important questions unanswered; for example, are both high and low expectancy groups equally successful in obtaining expected results or does only one of the two groups obtain results significantly different from the no-expectancy control group? Obviously the interpretation of the E-bias effect would be quite different depending upon which of these two processes actually occurred. A no-expectancy control group of Es was therefore included in the present study in order to shed further light on such questions.

METHOD

The method used in the present study follows Rosenthal as closely as possible with respect to development of the experimental task, selection of Ss and Es, and instructions administered to Ss and Es. A complete description of Rosenthal's procedures can be found in his recent book (Rosenthal, 1966).

The photo-rating (PR) test consisted of 15 photos of male faces having previously received mean ratings between +1 and -1 when rated on a scale from +10 (successful) to -10 (failure). Fifteen upper-division male students served as Es and each administered the PR test to three to six Ss who were drawn from introductory psychology classes.

Procedure

Each E was scheduled for a 2-h block of time. During the first 15 min, the author familiarized E with the ostensible purpose of the study (collecting normative data on a newly-developed test in relation to certain groups of Ss), and discussed such matters as test materials and procedures for obtaining Ss and administering the test. The only differences in the instructions given to Es were as follows:

For (+ or -) 5 Es

"As I already mentioned, you'll be collecting data from certain groups of Ss. The Ss you'll be running today were selected on the basis of questionnaire and inventory data collected at the beginning of the semester and they should average about a (+ or -) 5 rating."

For no-expectancy Es

"As I already mentioned, you will be collecting normative data for certain groups of Ss. The Ss you'll be running today will make up part of a control group. They were randomly selected from introductory psychology courses without any reference to the questionnaire and inventory data collected at the beginning of the semester."

At this point, Es were given a list of their Ss and the times for which they were scheduled. Finally, they were told that since Ss were introductory psychology students we would have to give them 1 h credit, and since the administration would only take a small part of an hour, we were going to make use of Ss in another study being conducted by another graduate student. Es were therefore asked to send Ss to another room after completion of the PR test.

The actual purpose of this procedure was to administer an "awareness" questionnaire to each S in an attempt to examine the relationship between the E-bias effect and S's awareness (or recognition) of E's expectancies. After E contacted his last S, he was also given a brief questionnaire in order to evaluate the relationship between E-bias and E awareness of the expectancy induction.

Experimental Design

A one-way analysis of variance was used to analyze the primary data regarding the E-bias effect. The single factor was comprised of three levels of E expectancy: one group of Es expecting +5 ratings, one group expecting -5 ratings, and a third group of Es with no-expectancy. Each of the three levels contained five Es, each administering the PR test to three to six Ss. The actual experimental unit for each E was the mean rating assigned to all 15 photos averaged over all Ss contacted by the E.

The last item on the subject questionnaire asked S to select that rating (from +10 to -10) he believed E expected to obtain. These ratings were then tested for significance—that is, in order to determine whether S's perception of E's expectancy was related to the different expectancy groups. Two t tests (for the three possible comparisons) were calculated in order to evaluate this question.

RESULTS

A one-way analysis of variance based on the mean photo-rating obtained by Es in the three expectancy groups yielded a nonsignificant value ($F < 1.0$). The mean ratings obtained by the three groups of Es were as follows: .70 for +5 Es, .94 for -5 Es, and 1.07 for nonexpectancy Es.

As noted in the previous section, Ss were administered an "awareness" questionnaire and were asked to state that rating which E expected, regardless of Ss' confidence in such a judgment. The mean expected ratings assigned by Ss to Es are as follows: .38 for +5 Es, 1.17 for -5 Es, and 3.04 for

no-expectancy Es. A comparison between the +5 and no-expectancy groups yielded a significant difference ($t = 2.58$, $df = 46$, $p < .02$, two-tailed) whereas the -5 group and no-expectancy group were not significantly different ($t = .57$, $df = 40$, $p < .50$, two-tailed). When the expected ratings assigned to Es by Ss were dichotomized (positive vs negative ratings), a chi-square test also yielded a significant value ($\chi^2 = 6.52$, $df = 2$, $p < .05$). Inspection of the other items on the S questionnaire (as well as the brief questionnaire given to Es) yielded no additional information.

DISCUSSION

The most important finding to emerge from this study was the inability to replicate the E-bias effect within the framework of Rosenthal's basic paradigm. Coupled with the Barber and Silver conclusions, the present finding indicates that assuming the stability of the effect and focusing on particular modes of mediation is rather premature, and it is suggested that subsequent research be directed toward specification of conditions under which the effect actually emerges.

Barber and Silver suggest a number of variables which appear related to the emergence of the E-bias effect; for example, the relationship between Es and the principle investigator, the methods used to induce bias, and the nature of the experimental task. All of these variables need much more attention. In the present study the instructor of the class from which Es were drawn was not the individual who administered instructions to Es, and there was no payment for Es' participation. In addition, a number of Es expressed suspicion about the stated purpose of the study during the post-experimental interview. The exact manner by which such conditions might have affected the emergence of an E-bias

effect is not yet clearly understood, and subsequent research activities of the author will be directed toward just such factors.

As noted in the previous section, analysis of the S questionnaire data revealed that no-expectancy Es were seen by Ss as expecting higher ratings than were the other groups of Es. In light of the inability to obtain an E-bias effect, this finding appears less meaningful and interpretable than might have been the case with a significant E-bias effect.

REFERENCES

- BARBER, T. X., CALVERLEY, D. S., FORGIONE, A., McPEAKE, J. D., CHAVES, J. R., & BOWEN, B. *Five attempts to replicate the experimenter bias effect*. Harding, Mass.: Medfield Foundation, 1966 (Mimeo).
- BARBER, T. X., & SILVER, M. J. Fact, fiction, and the experimenter bias effect. *Psychological Bulletin*, in press.
- FRIEDMAN, N. The psychological experiment as a social interaction. Unpublished doctoral dissertation, Harvard University, 1964.
- KINTZ, B. L., DELPRATO, D. J., METTEE, D. R., PERSONS, C. E., & SCHAPPE, R. H. The experimenter effect. *Psychological Bulletin*, 1965, 63, 223-232.
- MASLING, J. Role-related behavior of the subject and psychologist and its effects upon psychological data. In D. Levine (Ed.), *Nebraska symposium on motivation*. Lincoln: University of Nebraska Press, 1966. Pp. 67-103.
- McGUIGAN, F. J. The experimenter: A neglected stimulus object. *Psychological Bulletin*, 1963, 60, 421-428.
- ROSENTHAL, R. *Experimenter effects in behavioral research*. New York: Appleton-Century-Crofts, 1966.
- ROSENTHAL, R., FRIEDMAN, N., & KURLAND, D. Instruction-reading behavior of the experimenter as an unintended determinant of experimental results. *Journal of Experimental Research in Personality*, 1966, 1, 221-226.