

## EFFECTS OF SUCCESS AND FAILURE ON CONFIDENCE

M. FRANK NORMAN

*Dartmouth College*<sup>1</sup>

The present study is concerned with alterations in *S*'s confidence in his ability to perform a certain task correctly. These alterations were brought about by manipulation of *S*'s belief about the correctness of his previous performance of this task. A correct performance will be referred to as a *success*, an incorrect performance as a *failure*.

This study was undertaken to test a mathematical model for alterations in confidence in success-failure sequences. It was found that the model did not predict successfully, and it is of interest here only because one of its assumptions was of independent psychological interest and could be tested directly from the data. This assumption (Hypothesis I) was that the effectiveness of success in raising and of failure in lowering confidence is constant over all trials in the experimental situation to be described below.

It was also possible to test the following prediction on the basis of the data collected: Success more often leads to a rise than to a fall in confidence, while failure more often leads to a fall than to a rise in confidence (Hypothesis II). More precise statements of Hypotheses I and II will be given subsequently.

The changes in confidence which result from success and failure have, as far as the author has been able to discern, received extensive experimental attention only in two investigations (Child & Whiting, 1949; Bayton & Whyte, 1950). Both of these investigations were methodologically quite different from the present study, and there seems to be little basis for comparison. However, the problem of confidence changes after success and failure has an obvious formal similarity to the problem of changes in expected performance score in response to success and failure as studied in expectancy and level of aspiration experiments. Thus some of the findings in these areas are worth mentioning here. The effectiveness of success and failure in altering expected performance score has been found to decay with increasing trials (Steisel & Cohen, 1951; Castaneda, 1952). This effect was predicted by Rotter (1954, pp. 174-182). It also has been shown repeatedly that success tends to raise expected performance score while failure tends to lower it (Gruen, 1945; Ausubel & Schiff, 1955). In all of these studies success and failure were defined in a manner different from but consonant with the definition employed in this study.

### METHOD

Seventeen Psychology 1 students served as *Ss*. These men were motivated to participate by the offer of the addition of a few "points" to their overall course standing.

The experimental room was dimly illuminated and *S* sat close to and facing a wall, on which was a round piece of white cardboard 24 in. in diameter. *S* was provided with an unsharpened pencil. *E* sat at a desk, on which stood a pad of graph paper, 8 ft. behind and slightly to the right of *S*. The latter was instructed as follows:

"You are about to participate in a study of a certain type of problem-solving ability. The problem you will be asked to try to solve is, in many ways, an unconventional one, and scores on this test do *not* correlate very well with either standard aptitude or achievement test scores.

<sup>1</sup>Now a student in the Department of Statistics, Stanford University. The author is grateful to William M. Smith for his suggestions in connection with the preparation of this paper for publication.

"We will proceed as follows. Before you, you see a large round piece of white cardboard. If you glance over your right shoulder, you will see, on my desk, the back of a pad of graph paper. On this pad, facing me, is a symmetrical fine line pattern. Here are copies of some patterns which have been used in the past, though neither of these will necessarily be used in this experiment. [Sample patterns given to *S*.] It will be your task to try to discern what the pattern on my pad is in the following way: With the eraser end of the pencil on your desk, you will touch the cardboard in front of you. If you touch a spot on your cardboard such that one of the lines of the pattern crosses the comparable spot on my pad, I will say 'correct.' Otherwise I will say 'incorrect.' These corrections will help you modify your choices. [Sample patterns collected.]

"This is a very difficult type of problem to solve, and we do not believe that anyone can completely master it. Consequently scores on this test run lower than those on most tests with which you are familiar. However, those who possess the requisite ability, though they never quite understand what the pattern is, eventually come to point correctly fairly often. Those who do well on the test are never able to say much about their method of solving the problem—typically they insist that when they relax but concentrate hard they just seem to do well.

"You will receive one point for every correct choice you make—I will keep score. It is, incidentally, a violation of the rules to deliberately touch the same point twice. If you do, the second time you touch the point I will tell you 'correct' half the time and 'incorrect' half the time, regardless of what the pattern is at that point. This is to prevent you from running up a high score by touching the same correct point over and over again. It should be mentioned that you will be asked to make your choices fairly quickly.

"One final feature of the test is the following. Before each choice—before each time you 'point'—you must state how confident you are that your next choice will be correct. You will do this by simply saying 'high confidence,' meaning that you think you have a fairly good chance of being correct on your next choice, or 'low confidence,' meaning that you don't think you have much of a chance of being correct on your next choice. You must express your level of confidence as either 'high confidence' or 'low confidence.'

"To summarize: State your level of confidence, touch a spot on the cardboard, and I will tell you if you were correct or not. Then you will state your level of confidence in your next choice, make your next choice, and so on."

There was, in fact, no pattern on *E*'s pad, and the outcome ("correct" or "incorrect") of each trial ("choice") was determined before the start of the experiment by means of a table of random numbers. During each of 4 blocks of 20 trials, *S* was told that he had pointed correctly 9 times (except for one *S* who was told in his third block of 20 trials that he was correct 10 times and, in his fourth block of 20 trials, that he was correct 8 times). No three *S*s received the same random sequence of outcomes.

*S*s were allowed to proceed at their own pace provided that their response rates were between four and eight responses per minute. There was a 2-min. break for each *S* after his 50th trial. From all indications *S*s took the instructions at face value and thus regarded the task as, in some sense, skill dependent. The interest of almost all *S*s seemed high throughout the experiment.

## RESULTS

The unit of analysis for this study was the triple of responses consisting of a statement of confidence by *S*, a statement of correctness or incorrectness by *E*, and the next statement of confidence by *S*. For each *S*, each trial (except the

first) yielded one of eight distinct triples of this sort. Letting  $l$  stand for "low confidence,"  $h$  for "high confidence,"  $I$  for "incorrect," and  $C$  for "correct," any one triple of responses can be identified by a sequence of three letters, e.g.,  $lCb$ . Letting  $\Sigma lCb$  stand for the total number of  $lCb$  triples which appear in the block of trials under consideration,<sup>2</sup> we define, for any given  $S$  and block of trials, the quantities  $d$ ,  $e$ ,  $f$ , and  $g$  as follows:

$$d = \Sigma lCb / (\Sigma lCb + \Sigma lCl)$$

$$e = \Sigma hIl / (\Sigma hIl + \Sigma hIb)$$

$$f = \Sigma lIb / (\Sigma lIb + \Sigma lIl)$$

$$g = \Sigma bCl / (\Sigma bCl + \Sigma bCb)$$

For any given  $S$  and block of trials,  $d$  gives the ratio of the total number of times that confidence was raised by success to the total number of times that confidence might have been raised by success. The quantity  $d$  will be referred to as the relative frequency of confidence rise after success or as the effectiveness of success in raising confidence. Similarly,  $e$  will be referred to as the relative frequency of confidence fall after failure or as the effectiveness of failure in lowering confidence. The quantity  $f$  will be referred to as the relative frequency of confidence rise after failure, and the quantity  $g$ , as the relative frequency of confidence fall after success. It should be noted that the denominator in one of these expressions was occasionally zero, in which case the expression was undefined, and  $S$  having such an undefined ratio was necessarily omitted from that part of the analysis. The  $N$  for which a given result pertains is derivable from the  $df$  listed with the corresponding test in the paragraphs below.

In terms of the quantities defined above, Hypothesis I asserts the constancy of  $d$  and  $e$  over trials. To test this hypothesis both  $d$  and  $e$  were computed for each  $S$  for the first and for the second block of 40 trials. The effectiveness of success in raising confidence,  $d$ , showed a mean decline of .147 from the first to the second 40 trials. The corresponding  $t$  was 3.35, and thus this decline was significantly different from 0 at the .005 level with 15  $df$  in a two-tailed paired-replicates  $t$  test. The effectiveness of failure in lowering confidence,  $e$ , showed a mean decline of .079 ( $t = 1.25$ ) from the first to the second 40 trials. The probability of a decline as great as this under Hypothesis I was greater than .20 (16  $df$ ). (Similar comparisons reveal a significant mean increase in  $f$ , the relative frequency of confidence rise after failure [mean increase = .124,  $t = 2.62$ ,  $df = 14$ ,  $p < .05$ ], and a nearly significant mean decline in  $g$ , the relative frequency of confidence fall after success [mean decline = .134,  $t = 1.95$ ,  $df = 16$ ,  $.05 < p < .10$ ], from the first to the second 40 trials.)

Hypothesis II asserts that  $d > g$  and  $e > f$ . To test this hypothesis  $d$ ,  $e$ ,  $f$ , and  $g$  were computed for each  $S$  over all 80 trials. The mean difference of .41 between  $d$ , the relative frequency of confidence rise after success, and  $g$ , the relative frequency of confidence fall after success, was significantly greater than 0 at the .001 level ( $t = 3.98$ ,  $df = 16$ ) in a one-tailed paired-replicates  $t$  test. The mean difference of .20 between  $e$ , the relative frequency of confidence fall after failure, and  $f$ , the relative frequency of

<sup>2</sup>Table A, giving  $\Sigma lCb$ , etc., and also the total number of "high confidence" responses, for each  $S$  for the first and second blocks of 40 trials, has been deposited with the American Documentation Institute, Auxiliary Publications Project, Photoduplication Service, Library of Congress, Washington 25, D.C. Order Document No. 6880 remitting \$1.25 for 6- by 8-in. photocopy or 35-mm. microfilm.

confidence rise after failure, was significantly greater than 0 at the .05 level ( $t = 1.96$ ,  $df = 16$ ).

These results point to the conclusion that Hypothesis I is in error, at least in so far as it predicts no change in the effectiveness of success in raising confidence as a function of number of elapsed experimental trials. On the other hand, results suggest that Hypothesis II is correct. The degree to which these conclusions will be found applicable to other types of experimental situations is, of course, an open question.

Finally, a Pearson product-moment correlation of  $-.485$  between  $d$ , the effectiveness of success in raising confidence, and  $e$ , the effectiveness of failure in lowering confidence, computed over all 80 trials was obtained. This  $r$  was significantly different from 0 at the .05 level in a two-tailed test ( $N = 17$ ). Thus, for Ss for whom success was relatively effective in raising confidence, failure tended to be relatively ineffective in lowering confidence.

#### SUMMARY

Each of 17 Ss was given 80 trials on a task which involved an action of  $S$  which  $E$  labeled "correct" or "incorrect." Ss were led to believe that success or failure on each trial was skill determined, while in reality successes and failures occurred in accordance with a predetermined random sequence containing a fixed number of each. Before each trial, each  $S$  characterized his confidence in his forthcoming performance as high or low. (a) The effectiveness of success in raising confidence declined from the first to the second 40 trials. The relative frequency of confidence rise after failure increased from the first to the second 40 trials. (b) The relative frequency of confidence rise after success was greater than the relative frequency of confidence fall after success. The relative frequency of confidence fall after failure was greater than the relative frequency of confidence rise after failure. (c) The effectiveness of success in raising confidence and the effectiveness of failure in lowering confidence were negatively correlated.

#### REFERENCES

- AUSUBEL, D. P., & SCHIFF, H. M. A level of aspiration approach to the measurement of goal tenacity. *J. gen. Psychol.*, 1955, 52, 97-110.
- BAYTON, J. A., & WHYTE, E. C. Personality dynamics during success-failure sequences. *J. abnorm. soc. Psychol.*, 1950, 45, 583-591.
- CASTANEDA, A. A systematic investigation of the concept expectancy as conceived within Rotter's social learning theory of personality. Unpublished doctoral dissertation, Ohio State Univer., 1952.
- CHILD, I. L., & WHITING, J. W. M. Determinants of level of aspiration: evidence from everyday life. *J. abnorm. soc. Psychol.*, 1949, 44, 303-314.
- GRUEN, E. W. Level of aspiration in relation to personality factors in adolescents. *Child Developm.*, 1945, 16, 181-188.
- ROTTER, J. B. *Social learning and clinical psychology*. Englewood Cliffs, N. J.: Prentice-Hall, 1954.
- STEISEL, I. M., & COHEN, B. D. The effects of two degrees of failure on level of aspiration and performance. *J. abnorm. soc. Psychol.*, 1951, 46, 79-82.

Accepted September 9, 1961.