

THE CONTROL OF THE CONTENT OF CONVERSATION THROUGH REINFORCEMENT

N. H. AZRIN, W. HOLZ, R. ULRICH, and I. GOLDIAMOND¹

ANNA STATE HOSPITAL, ILLINOIS²

The present study is an attempt to replicate an investigation by Verplanck (1955). In that investigation, psychology students were reported as able to exert strong control over the casual conversations of other people by selectively reinforcing a certain type of opinion-statement and extinguishing all other types of statements. To produce reinforcement, the student *E* paraphrased or agreed with each opinion of the *S*. Conversely, to produce extinction, the student *E* openly disagreed with the opinions or simply said nothing. Each *S* was studied for 30 minutes. For some groups of *S*s, 10 minutes of extinction was preceded and followed by 10 minutes of reinforcement. For other groups, a 10-minute period of reinforcement was preceded and followed by 10 minutes of extinction.

The results of this experiment were quite dramatic. Within this brief 30-minute period, "All *S*s increased their rate of stating opinions, regardless of the topic of conversation, its setting, or *S*'s particular relationship with the *E*." (Verplanck, 1955, p. 673).

From the theoretical view, this study is one of the very few successful attempts to reinforce human verbal behavior in a free-operant situation. From a more practical view, this study seemed ideal as a laboratory exercise for a class of college graduate students. No special laboratory apparatus was needed other than a clock or watch which permitted the students to record the number of opinions and statements at 1-minute intervals. No laboratory space was needed because the experiment had been conducted in informal settings, such as a cafeteria or dormitory, or even over a telephone. Other advantages were the brief period of time required (30 minutes) and the apparent simplicity of the procedure. For example, after adequate instruction, "Of the 17 students who undertook the experiment, all were able to collect one or two sets of data as the design demanded." (Verplanck, 1955, p. 669).

CLASS EXPERIMENT I

Unlike the original study, a class composed of graduate students rather than undergraduates was used here. Of the 16 students, 11 either majored or minored in psychology. Before replicating this experiment, these students had received intensive instruction in the principles of operant conditioning, and had been tested on their knowledge. Further, each student read Verplanck's original study and was examined on his knowledge of it, and the procedure was discussed and rehearsed in detail in class. Two procedures were used. Half of the students conducted 10 minutes of extinction followed by 10 minutes of reinforcement, which was followed by another 10 minutes of extinction. The other half of the students reversed the procedure: reinforcement—extinction—reinforcement. Reinforcement consisted of the student *E*'s agreement with the opinions of the subject. Extinction consisted of saying nothing. As in the original study, the student *E*s were instructed to say nothing at any time other than when agreeing with expressed opinions.

The results here were quite similar to those of the original study. Out of 15 students, 14 reported a higher frequency of opinions during the reinforcement period than during

¹Now at Arizona State College, Tempe, Arizona.

²This investigation was supported by a grant from the Psychiatric Training and Research Fund of the Illinois Department of Public Welfare.

extinction. An application of the test by Wilcoxin for paired replicates³ to these results showed that they were statistically significant well below the .0001 level. As was also true of the original study, the frequency of statements other than opinions did not change greatly.

Except for one student who could not complete the experiment, none of the students reported any difficulty either in conducting the experiment, defining the responses, scoring the behavior, or obtaining the expected results. However, the one student reported that he was unable to maintain a conversation without actively participating in the conversation, although such participation clearly violated the prescribed procedure. Also, he found it difficult to categorize the conversation into opinion vs. statements while simultaneously attending to his watch for purposes of timing. After six attempts, this one student reported that he was unable to complete the experiment according to the required procedure. After class discussion of his difficulties, eight of the other students stated that they had had similar difficulties, previously unmentioned, and were thereby forced to deviate appreciably from the stated procedure. The most frequently stated difficulties were: (1) the *S* walked out during extinction; (2) the *E* was forced to actively participate by nodding, smiling, or asking questions; (3) the *S* was aware of the recording; (4) *E* made errors in timing; and, finally, (5) *E* became too interested in the subject of conversation to concentrate on the recording of opinions.

During this class discussion, one student mentioned that he had taken a tape recording of his own experiment. A second student was therefore assigned to analyze the tape to determine the reliability of the recording. After several attempts, this second observer reported that he could not discriminate between the reinforcement and extinction periods, since the *E* had been actively participating throughout. A comparison of the frequency of opinions and statements between the two observers revealed little or no correspondence.

Because of this apparent difficulty in categorizing verbal behavior as an opinion or a statement, the following study was conducted.⁴ A 30-minute tape recording was taken of a conversation in which the speaker knew that his opinions and statements were being recorded. Five students from this same class, including the original *E*, then listened to the recording individually and recorded the numbers of opinions and statements made. To minimize errors in timing, the five observers noted the behavior in terms of 5-minute rather than 1-minute intervals.

Figure 1 presents the number of opinions recorded by the five observers. It shows that during each 5-minute interval, the number of opinions differed greatly among observers. The minimum and maximum is in a ratio of about 2 to 1 for each 5-minute period. More surprising is the difference found between the *E*'s original scoring and his own later scoring from the tape recording only 2 hours later. (See dark lines in Fig. 1.)

CLASS EXPERIMENT II

The many difficulties reported by the first class in replicating the original experiment, as well as the surprising lack of reliability in identifying the responses, might be interpreted simply as differences in the types of students used. A second class was therefore given the same experiment, but was not given the original article to read. Furthermore, they were not told that the study had ever been done before. As in the original study by Verplanck, this class was composed largely of undergraduates who had been instructed and tested on the

³This test was used in all statistical comparisons in this paper.

⁴Mr. R. Teague assisted in the collection of the data for this aspect of the study.

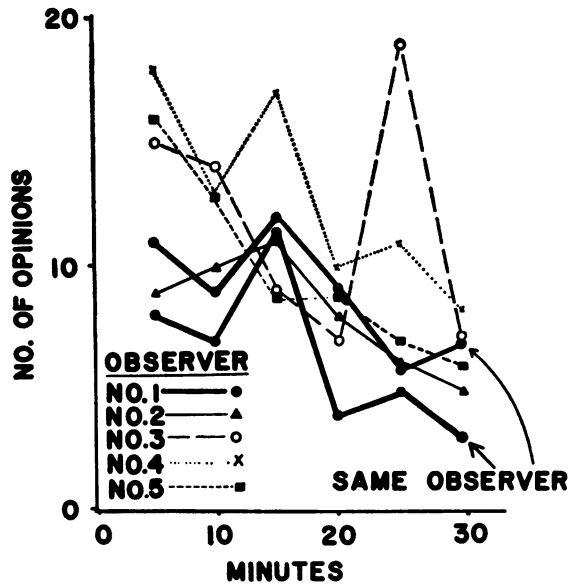


Figure 1. Number of opinions identified by each of five observers listening to the same tape. Observer No. 1 (heavy lines) listened to the conversation a second time after an interval of 2 hours.

principles of reinforcement. The procedure followed for this second class was the same as that followed in the previous class. There were successive 10-minute periods of reinforcement—extinction—reinforcement. However, extinction was defined as disagreement rather than as silence. Verplanck had found that disagreement was as effective as silence in decreasing the response, and the present results support these findings. Of the 12 students who completed the study, 11 found that disagreement (described as extinction) produced a lower frequency of opinions than did agreement (described as reinforcement). This difference was again statistically significant at below the .001 level. Again, the students reported no serious difficulty.

These results pose somewhat of a dilemma. Despite seemingly great unreliability in identifying an opinion, and despite the numerous difficulties the first class of students mentioned, agreement with opinions is reported almost unanimously to produce a higher frequency of opinions that does silence or disagreement with the opinions. This procedure also seems to work as well for undergraduates as for graduate students.

CLASS EXPERIMENT III

An analysis of the results of the previous classes revealed some apparent relationship between the student's understanding of reinforcement principles and his reported results. For example, those students who were frequently absent from class and who revealed relatively little knowledge of reinforcement on their class exams usually reported relatively little effect of reinforcement upon opinions. At least two reasons might account for this failure: either these students did not follow the appropriate procedure, or they had no firm expectations about the results. To evaluate this second possibility, a third class of undergraduates was assigned to the same experiment, but this class had not read the original

article. Half of the students conducted consecutive 10-minute periods of reinforcement—extinction—reinforcement, and the other half conducted periods of extinction—reinforcement—extinction. Again, reinforcement consisted of agreeing with each opinion, whereas extinction consisted of silence. This experiment was conducted during and after classroom instruction on the principles of reinforcement. The results were virtually the same as those in the previous class: 44 out of 47 students reported a greater frequency of opinions during reinforcement than during extinction ($P < .001$). One month later, following classroom discussion of emotions, essentially the same experiment was assigned to these students except that periods of disagreement were used instead of periods of silence. Verplanck had found that silence and disagreement were functionally identical in producing a decrease of opinions. However, the effect of disagreement was discussed in the present class in terms of catharsis rather than in terms of extinction, as in Verplanck's class and in the second class of the present experiment. The students were told that catharsis is the release of emotion, and that such release could be obtained by agreeing with an individual, so that a relative state of tranquility followed in which there was little reason for stating strong opinions. Conversely, disagreement was described as preventing catharsis, so that an increase of emotionally charged opinions followed. These comments about catharsis were intended to produce a bias toward obtaining an increase in the frequency of opinions rather than the decrease that was reported by Verplanck and that was obtained in the previous class in the present study.

Figure 2 compares the mean number of opinions reported by the students of this third class with the number reported by the students of the second class under the two sets of response bias. According to Fig. 2, the procedure of disagreement produced an increase of opinions when identified as catharsis but a decrease when identified as extinction. Statistically, the difference in the number of opinions between agreement and disagreement for

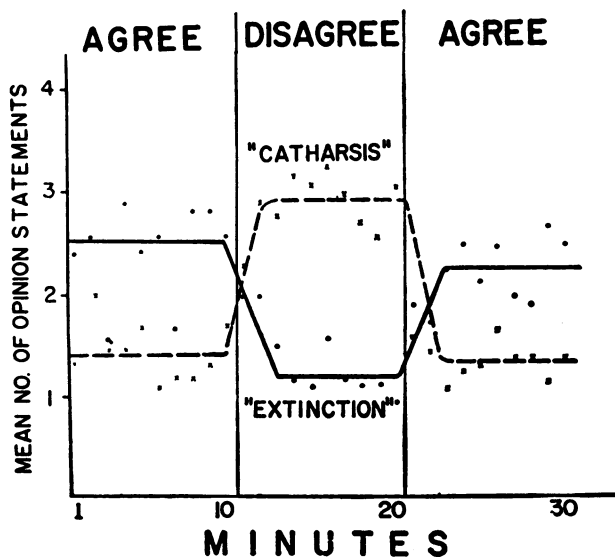


Figure 2. The effect of the expectations of experimenters upon the results obtained. One group of *Es* (dotted line) was told that disagreement was expected to produce an increase because of catharsis. The other group of *Es* (solid line) was told that disagreement (the same procedure) would produce a decrease of opinion.

each class, as well as the difference between the two disagreement periods, is significant at well below the .001 level. Allegedly, the procedure of disagreeing with the subjects' opinions was the same for both classes. Therefore, the actual procedures used appeared to be almost irrelevant for obtaining the above changes in opinions. Rather, the same procedures of disagreeing with opinions have resulted in diametrically opposite effects because of the response bias given to the experimenters.

It will be recalled that the first class of graduate students had reluctantly admitted to various procedural difficulties. By coincidence, a student was enrolled in this third class who was also employed as a research assistant in a psychology laboratory. This student employee was assigned to question the other students informally and outside of class as to how they had conducted their experiments. The other students had no knowledge of this arrangement. Out of 19 students questioned, consisting of almost one-half of the class, 12 stated that they fabricated part or all of the data. This admission of "dry-running" was readily made when the student was asked by the employee, "I'm having trouble with my experiment; can you tell me how you did yours?" Five of the remaining seven students questioned stated that they had deviated greatly from the prescribed procedure. Only two out of nineteen students stated that they had followed the prescribed procedure. Consequently, an attempt at an exact replication seemed pointless, since the data reports themselves were probably fabricated.

Granted that students may not be competent investigators, the question still remained whether or not more experienced investigators could demonstrate the control of casual conversation by reinforcement. The study was therefore repeated by four experienced investigators. Each of these four investigators had (1) advanced psychology training—a Ph.D. or M.A.; (2) skill in shaping animal behavior; (3) a healthy respect for negative results; and (4) practice in defining opinions. One might reasonably and justifiably consider these four criteria as essential in guaranteeing the success of the reinforcement procedure. In order to maximize the likelihood of success, a sequence of agreement—disagreement—agreement was used to avoid the expected difficulties in maintaining silence. Out of 12 attempts, not one of the four *E*'s could complete his experiment. It may be recalled that the procedure requires that the *E* restrict himself to agreement (or disagreement) of opinions, and stipulates no questions, statements, nods, smiles, or other types of interaction. The reason for forbidding such behavior proved to be obvious: *E*'s reaction, however subtle, could often be seen to exert profound but uncontrolled effects upon the conversation of the subject. In the absence of any reaction by the four *Es*, however, all of the twelve *Ss* terminated the conversation within 10 minutes by leaving the room where the conversation was taking place.

The above results indicate that the successful reinforcement of opinions by student experimenters in casual conversation seems open to serious question. Out of a sincere attempt to produce such reinforcement, procedural difficulties emerged, also unreliability in identifying the response, results that were a function of *E*'s expectations, and, finally, evidence of extensive falsification of the data. These findings have implications for the conduct of research in operant conditioning. Operant-conditioning procedures have generally been characterized by a high degree of control. In order to avoid unreliability, the response is usually defined very simply and precisely. In order to ensure proper programming of the procedure, automatic apparatus is used. Printed records of the responses are also obtained by automatic means to eliminate bias from the *E*'s expectations. These and other precautions have been used, not because of any inherent fascination with "artificial" situations or with com-

plex equipment, but because empirical considerations have demanded such control. The importance of extending the procedures of operant conditioning to "real-life" situations should not be allowed to override the elementary considerations of experimental control.

It would be incorrect to conclude from this study that control over verbal behavior through reinforcement is not possible. Indeed, the existence of such control seems demonstrated by the degree to which the reports of the experimenters could be modified. Certainly, if the situation could be structured in such a manner that the *E* is not expected or required to speak, there would seem to be a reasonable basis for selective reinforcement. However, the present findings indicate that different procedures will be needed before free-operant conditioning can be validly extended to the control of casual conversation. More important, these findings emphasize the necessity of objective programming and recording in the study of verbal conditioning. In the absence of such objectivity, the results of studies in verbal conditioning may be more of a reflection of the experimenter's expectations and theories than of the subject's behavior.

REFERENCE

Verplanck, W. S. The control of the content of conversation: reinforcement of statements of opinion. *J. abnorm. soc. Psychol.*, 1955, **55**, 668-676.

Received July 1, 1960.

This article, reprinted by permission, originally appeared in the *Journal of the Experimental Analysis of Behavior*, 1961, **4**, 25-30.