

# Transfer of rewarded responses in personality judgments

MELVIN H. MARX

*University of Missouri, Columbia, Missouri 65201*

A total of 108 students in five high-school psychology classes judged 32 facial photographs on four personality traits. Students worked in pairs, one making the judgments ("performing"), the other checking correct answers ("observing") when the experimenter called them out after all subjects had judged the face on all four traits (emotionality, intelligence, shyness, and sincerity). Students traded roles after the first 16 items. Male performers selected previously correct traits more often than did male observers when they were asked to select the "most outstanding" trait for each face in a subsequent test. This result provides some support for the predicted strengthening by reward with information controlled. No difference occurred for females.

This experiment is one in a series designed to determine whether reward can be shown to have response-strengthening effects when the informational content of feedback is controlled. Recent studies have shown that subjects who perform (make behavioral decisions and receive feedback) generally tend to repeat both their correct and their incorrect responses more often than matched observers who receive equivalent feedback information (Marx & Witter, 1972; Marx, Witter, & Farby, 1973; Witter, Marx, and Farby, 1976).

In the first experimental manipulation of the perform/observe variable, Hillix and Marx (1960), utilizing a transfer test, found some evidence that performers tended more than observers to return to old learned responses when attempting to solve new problems. Marx and Witter (1972) later found that performers were superior in recognizing full faces after having learned to identify parts of those faces.

The present experiment represents an attempt to replicate these results in a situation where both training and transfer were arranged as tests of the ability to judge personality traits from facial photographs. The task was utilized in an effort to maintain a relatively high level of subject interest and to make the transfer test relatively unobtrusive, so as to elicit more natural behavior for the experimental analysis.

Thanks are due Ms. Joan Girnis and Mr. Gary Blodick, teachers of the psychology course at Naples High School, Naples, Florida, for their cooperation in making this experimentation a part of the course work in an effort to familiarize their students with "real-life" psychology by giving them some experience as experimental subjects. In class periods after the collection of the data, an explanation of the design and purpose of the experiment was made, some preliminary results were presented, and questions were answered by the experimenter. This research was supported in part by Research Career Award 5-K06-MH-22023 from the National Institute of Mental Health and a grant from the U.S. Army Research Institute for the Social and Behavioral Sciences. The opinions expressed herein are those of the author and in no way are to be considered as endorsed by the U.S. Army.

## METHOD

### Subjects

The subjects were 63 female and 45 male high-school students, trained and tested in their normal classroom situation. They were students in five psychology classes, ranging in size from 18 to 30 in attendance on the day of the experiment.

### Design

Students were placed in like-sex pairs as they entered the classroom. Within each pair, one student served first as a performer for half of the training items, then as an observer for the other half; the other student served first as an observer, then as a performer. Correct answers were provided following all the performers' judgments on each face. The pairs were then separated for the second (transfer) phase, and the students were tested without feedback.

### Task and Procedure

The subject's task in this experiment was to judge personality traits from facial photographs. In preliminary research (Conover, Seymour, Marx, & Moore, Note 1), college students were asked to estimate the frequency with which each of eight personality traits had been applied to a large number of faces (college student photographs from old yearbooks). The four personality traits (emotionality, intelligence, shyness, and sincerity) that produced the greatest heterogeneity of variance for 32 representative faces (16 of each sex) were selected for use in this experiment.

The students used their ID numbers or their initials rather than their names on record sheets to minimize personal identification.

In the first (training) phase of the experiment, the performers' task was the same as the task used in the preliminary research mentioned above. They were asked to estimate the frequency with which each trait was likely to be applied to each of the faces shown. They used a 10-point scale (0 to 9). Low numbers meant low-frequency estimates (e.g., 0 indicated that the trait rarely or never would be applied) and high numbers meant high-frequency estimates (e.g., 9 meant that the trait always or almost always would be applied). Each face was shown for 20 sec to permit the four estimates.

After the four ratings for each face had been completed by all of the performers, the "actual ratings" were provided. The observers were instructed to circle any performers' estimate that was exactly the same as, or within one point of, the actual estimate given. It was assumed that this act would serve as a kind of reinforcement of the face-trait associations made by the

performers. Each set of 16 faces was half male and half female, randomly mixed within the set.

Observers were also told that they should check the estimates and the outcomes so that we could see how such practice affected their own scores. The subjects were not told that there would be another task. The "actual ratings" were modified slightly from the means obtained in the preliminary work (Conover et al., Note 1). This was done to insure that, in each set of four personality ratings for a given face, there would be one low score, two intermediate scores, and one high score. However, the relative positions of the various personality traits were maintained.

In the second (transfer) phase, which followed a brief rest interval, the subjects were told they would be shown the same faces again but were to make a "somewhat different judgment" on them. These judgments were done individually, and no feedback was provided. The subjects were told to "Imagine that you are involved in making a movie and that you have the job of making decisions about casting—or selecting—certain persons for different parts of the movie. Suppose that you are to screen a large number of candidates on the basis, first, of their photographs alone. Also suppose that your primary task is to rate each face in terms of personality traits. We now want you to rate each of the faces in this way, in terms of what you see as the *most characteristic* personality attribute of the four traits just used—emotionality, intelligence, shyness, and sincerity."

Subjects were given 10 sec to study each face and to write down one personality trait for each one. They were further instructed to make a global judgment.

## RESULTS

The numbers of previously reinforced traits selected as the most outstanding characteristic in the transfer test were almost identical for the performance and observation conditions: 378 for the female performers, 375 for the paired female observers, out of a total of 1,531 reinforced traits; 251 for the male performers, 249 for the paired male observers, out of a total of 829 reinforced traits. Individual classes were evenly divided with respect to this measure, with two showing a slightly higher performance mean and two a slightly higher observer mean for both male and female subjects (and in each case one class having identical means).

These essentially identical results for the perform/observe variable obviously required no statistical test and offered little encouragement that the hypothesized difference between performance and observation in transfer of reinforced responses would be revealed in other comparisons. Nevertheless, total scores are relatively gross measures, and a more analytic comparison was performed. In this analysis only those items (faces) were used in which (1) there was a single reinforced trait and (2) the single reinforced trait was not the one that had been assigned the highest score, of the four traits, for that particular face. The analysis consisted of determining the proportion of choices of the reinforced trait and of the highest score trait for performance vs. observation by sex. The rationale underlying this analysis was that on an a priori basis one would expect subjects to show some bias toward the traits that had been assigned the highest scores, because

a similar question was being asked in the transfer test (namely, which is the most outstanding trait?). The presumption that the trait most often attributed to a given face (indicated by the high training score) might also be the "most outstanding" trait seems reasonable. Thus, pitting reinforcement—which in this analysis always involved a trait with one of the lower frequency scores—against high-frequency scores was seen to be a direct and relatively rigorous means of assessing the transfer effects of reinforcement, assumed to be more effective in the performer than the observer process.

The results obtained in this analysis were more encouraging for the hypothesis under consideration. The total numbers of reinforced and high-score trait selections were computed. Overall, female subjects selected 62 reinforced as against 109 high-score traits in the performance and 71 reinforced as against 85 high-score traits in the observation condition. Male subjects selected 55 reinforced traits as against 52 high-score traits in the performance and 51 reinforced as against 62 high-score traits in the observation condition.

Because of the technical difficulties involved in statistical testing of these group data, with various subjects differentially contributing to the totals, subjects were categorized in accordance with the type of trait each had most often selected in the transfer test to permit a chi-square analysis of the data. They were classified as "reinforcers" or "nonreinforcers" (i.e., selectors of traits with high scores). Subjects with an equal number of each type of selection were excluded from this analysis. The results are shown in Figure 1. It is clear that only male performers were mostly "reinforcers"; male observers and female performers and observers all predominantly selected high-score traits.

Chi-square analyses of the frequency data shown in Figure 1 indicate that there was a reliable sex difference for performance ( $\chi^2 = 7.34$ ,  $p < .01$  for 1 df) but not for observation ( $\chi^2 = .07$ ).

## DISCUSSION

The interaction of sex and perform/observe variables indicated in Figure 1 and confirmed by the chi-square analysis offers some support for the hypothesis that reinforcement has effects that cannot be readily attributed to the straightforward cognitive processing of information provided by feedback signals. The fact that this result occurred only for male subjects obviously requires some restriction of its generality. However, it is consistent with some sex differences previously reported in the perform/observe experimental paradigm, for example, that males repeat more errors when performing than when observing (Marx, Witter, & Farby, 1973) and that males are generally superior as performers, both in repeating correct responses and changing errors in a binary-choice task; females are superior in these respects as observers (Marx & Conover, Note 2). A definitive theoretical interpretation of these and other sex differences found in recent perform/observe research awaits further research.

Another restriction on the generality of the transfer effect is the fact that it was not found in the comparison of total numbers of reinforced selections under performance and observation.

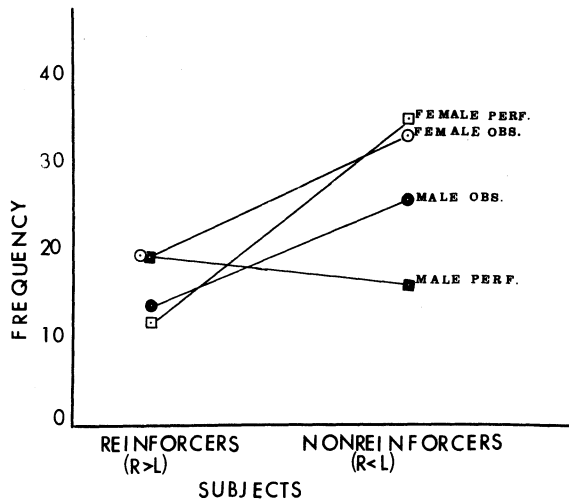


Figure 1. Categorization of male and female performers and observers as reinforcers (more reinforced responses transferred) or nonreinforcers (fewer reinforced responses transferred).

The failure of a gross measure to show a perform/observe difference, even while a more analytic measure does, has been found earlier; in that instance, performers and observers did not differ in number of correct responses, but performers repeated both correct and incorrect responses more often than observers (Marx & Witter, 1972). The methodological generalization suggested by these results is that it might well pay researchers to look beyond the standard, typically gross behavioral measures customarily utilized and to set up more specifically planned experimental confrontations in the data.

In spite of the qualifications, the fact remains that in this

experiment male performers at least did show a relatively greater tendency to transfer reinforced trait selections, thus offering further support for the hypothesis that reward can have behavioral effects that seem to be independent of the information that it provides.

#### REFERENCE NOTES

1. Conover, J. N., Seymour, G. E., Marx, M. H., & Moore, M. *Stereotyping of personality trait ratings based on facial photographs or subjects' prototypes*. In preparation.
2. Marx, M. H., & Conover, J. N. *Reward and information in binary-choice selective learning*. In preparation.

#### REFERENCES

- HILLIX, W. A., & MARX, M. H. Response strengthening by information and effect in human learning. *Journal of Experimental Psychology*, 1960, **60**, 97-102.
- MARX, M. H., & WITTER, D. W. Repetition of correct responses and errors as a function of performance with reward or information. *Journal of Experimental Psychology*, 1972, **92**, 53-58.
- MARX, M. H., WITTER, D. W., & FARBRY, J. Greater repetition of errors under performance compared to observation in multiple-choice human learning. *Perceptual and Motor Skills*, 1973, **37**, 949-950.
- WITTER, D. W., MARX, M. H., & FARBRY, J. Long-term persistence of response-repetition tendencies based on performance or observation. *Bulletin of the Psychonomic Society*, 1976, **8**, 65-67.

(Received for publication November 15, 1977.)