

Evaluating Industrial Trainers¹

ANDREW DuBRIN,² H. H. REMMERS and HARRY S. BELMAN

Recognition of the need for a valid and objective means of determining the effectiveness of individual trainers led to the development of an evaluation instrument called the Trainer Performance Indicator (T.P.I.). This was done by Richard Lanman,³ a graduate student, in a doctoral dissertation in cooperation with Dr. H. H. Remmers, Director, Division of Educational Reference, and Professor Harry S. Belman, Chairman, Industrial Education Curriculum, Purdue University. In this project and in subsequent experimentation with the TPI, valuable assistance was given by many training persons and a large num-

ber of business and industrial organizations. The study reported on here is a "cross validation" of the TPI. The scale which was administered to eight different groups in the spring and summer of 1957 demonstrated that it shows differentiations among trainers.

NATURE OF THE DEVICE

A forced-choice rating scale is one in which the rater is confronted with two or more statements, all of which seem to be equally favorable or unfavorable. The rater is then instructed to select one or two of these statements which best or

¹ Based on the Master's thesis by Andrew DuBrin under the supervision of H. H. Remmers.

² Now in the Department of Psychology, Michigan State University.

³ Lanman, Richard, "The Construction of a Forced-Choice Test for Industrial Trainers," unpublished doctoral dissertation, Purdue University, 1953.

HARRY S. BELMAN is Chairman of the Industrial Education Curriculum, Division of Education, Purdue University. Undergraduate and graduate work was done mainly at the University of Wisconsin. He has been an administrator of vocational education, has had industrial training experience, and during World War II served as Chief of Training, Region VI, War Manpower Commission. Since joining the staff of Purdue University he has carried on considerable research and other activity in the field of training. At present he is Chairman, Professional Standards Committee, ASTD.

H. H. REMMERS is Director of the Division of Educational Reference and Professor of Psychology and Education, Purdue University. An Iowa Ph.D., he is author and co-author of a number of books and has published upwards of 300 articles and monographs, chiefly in measurement and evaluation.

least describe the person in question. The scale constructor has available beforehand two statistical indices about each statement, which make scoring of such a scale possible: a favorability and a discrimination index.

THE favorability index simply tells whether the statements in each group appear to be equally favorable or unfavorable to most people. The discrimination index tells whether the statement in question is characteristic of good or poor trainers. For each block of statements the trainer gets a plus score if a statement characteristic of a good trainer is described as "most like him," or a statement characteristic of a poor trainer is described as "least like him." These indices are statistically determined by experimentation and preliminary try-out.

The TPI, consisting of three forms, was developed by this method. The study reported on here used one form of 65 blocks of two statements each. Raters checked one statement in each block as "most" or "least" like the trainer in question. Here is a typical item:

- Hovers over learners while they are practicing. (L)
- Has some irritating habits that bother the trainees. (L)

To determine whether the scale was actually valid (differentiated good trainers from poor trainers), sixty-three trainers representing eight different organizations were rated by a total of 622 trainees. It was necessary, of course, to establish some criterion measure. The criterion in every instance was a top-to-bottom ranking given the trainers by a training director or a superior who was

familiar with the work of all of the trainers in a given organization. Generally one trainer was rated by about ten trainees, using the TPI, and ranked among the other trainers by the training director in charge.

If the TPI is valid the scores on the forced-choice scale should be correlated positively with the rankings given by the training director; i.e., those men receiving high scores on the TPI should be ranked highly by the training director or superior. When the data for the eight organizations were combined the scores on the TPI correlated .54 with the criterion measures, a result that could have happened by chance less than one time in one hundred.

AN over-all validity of .54 for the TPI does not mean that the same results can be expected for every small group of trainers. This validity refers to the general effectiveness *across* organizations. We can expect considerable fluctuations when only a small number of trainers and trainees are used. Here is how the TPI worked out within the eight different groups.

Sixteen apprentice instructors (trainers) were rated by a total of 159 apprentices (trainees) or approximately ten to each trainer at the Arsenal Technical School in Indianapolis. The validity of the combined ratings was .33. The A. O. Smith Corporation of Milwaukee, Wisconsin and the Naval Avionics Facility of Indianapolis each submitted ratings for seven trainers, each group rated by 45 trainees. Here the correlations of the TPI with the training director's rankings were .49 and .46 respectively. Two Air Force ROTC training groups on the

Purdue campus rated nine instructors and the validities were .96 and .82 for groups of four and five instructors respectively.

The Commonwealth Edison Company of Chicago and the U. S. Post Office of Indianapolis each submitted data for six trainers who were rated by separate groups of trainees. The validities arrived at were .53 and .73 respectively. The largest individual group participating in the study was the plant of the IBM Corporation at Endicott, New York, where 13 trainers were rated by 216 trainees. In this case the validity was a substantial .73, statistically highly significant.

CHARACTERISTICS OF THE EVALUATION DEVICE

An essential requirement of any evaluating device is that it not only be valid, but it must also be reliable or consistent. It was necessary to establish not only how reliable the scale was itself but how reliable the training directors were in ranking the trainers. To determine the reliability of the scale itself, a statistical procedure ("split-half") was used which measures whether scores given on various parts of the scale correlate with scores given on parts of the scale. The reliability using this method was .79 and validity .54 across all trainers and raters.

Each training director who ranked the trainers was asked to rank them again, ten days later. Using this method it was

possible to establish whether the training directors' opinions of the relative merit of the men were stable, or fluctuated from week to week. Possibly because of the ease in remembering the ranking of ten names or less over a short period of time, the agreement was almost perfect in all cases.

THE conclusion drawn from the study is that the TPI is a usefully reliable and valid scale for evaluating industrial trainers, but at least two limitations should be mentioned. One limitation of the TPI in a training situation is its unsuitability for counselling. The training director cannot reveal specifically to the trainer where his weaknesses lie.

The TPI becomes increasingly reliable and valid when large numbers of trainers are each rated by a number of trainees; the more raters, the higher will be the reliability of their average rating. In general, less than ten trainees' ratings of a single trainer will be undesirably low in reliability.

The most important finding is that the instrument is reliable and valid across organizations necessarily varying widely in the content of their training.

Further experimental work must be done to establish norms which may serve as substitutes for the criterion established by the rank order list that a training director or superior may draw up for a group of trainers. The authors will be happy to arrange for cooperative use of the device by any organization.⁴

⁴ Address inquiries to Professor Harry S. Belman, Division of Education, Purdue University, Lafayette, Indiana.