

## A CRITICAL REVIEW OF TECHNIQUES FOR EVALUATING GUIDANCE

ROBERT M. W. TRAVERS

University of Michigan

THE term guidance covers numerous social activities in which one or more persons learn new behaviors through the help of another who is usually an adult. Guidance is essentially a learning situation and, in a sense, all teaching is guidance. In general, teachers and educators have come to draw little distinction between situations in which the pupil learns to add up sums of money, to make a speech, to organize a committee, to discuss current affairs, to choose a vocational goal, or to set up life goals. All of these are learning situations in which the function of the teacher is one of guidance; that is to say, one which facilitates learning. However, the term guidance is used by psychologists and by certain specialized workers in a different and limited sense to denote a limited range of learning situations in which a person with special psychological training plays an important part. Usually, in these learning situations the guidance worker is attempting to help the student to learn new behaviors which may solve immediate social difficulties or establish new long-term goals. The approach which the guidance worker takes is commonly called the *clinical* approach but the same approach when it is taken by the educator is referred to as the organismic approach. Both of these terms imply that the individual must be understood as a whole in order to understand his special problems of learning.

Counseling is a limited aspect of guidance and is a learning situation conducted on a personal basis by the guidance worker. Counseling is a learning situation even when it involves such complex phenomena as catharsis or transference. It is useful to think of it as a learning situation because it is then not only properly classified but it is also placed in a field in which there are well-developed evaluation techniques.

It should be noted that the teacher and specialized guidance worker agree that guidance refers to a learning situation, but the psychologist restricts the meaning of the term to those learning situations conducted by specialized guidance personnel. This paper is concerned mainly with the evaluation of guidance in this restricted sense. For the purposes of this discussion, it is irrelevant whether counseling is conducted on a non-directive or a "non-nondirective" basis. It is sufficient to recognize that both counseling and guidance represent attempts to facilitate learning and in this respect do not differ from other teaching situations.

In organized learning situations in education, goals are established, procedures are developed for attaining those goals, and methods are devised for determining the extent to which the goals are achieved. The latter process is now referred to as evaluation since it is used to determine the values inherent in the learning situation. During the last two decades great strides have been made in evaluating the outcomes of education and it is not uncommon to find schools in which serious attempts are made to measure not only the traditional subject-matter outcomes but also outcomes such as adequate social development, the appropriateness of the individual's leisure-time interests, and the adequacy of his vocational goals. However, the guidance movement, in the restricted and specialized sense of the term, has been largely uninfluenced by the evaluation movement partly because guidance workers commonly do not recognize that they, like any teacher, are trying to produce learning and therefore should measure how much learning has taken place, and partly because the evaluation of specialized guidance functions present special problems which will be considered here.

### *General Techniques for Evaluating Guidance*

There is no essential difference between the procedure for evaluating guidance and the procedure for evaluating the outcomes of any other learning situation provided by the school. Just as there are two general methods of evaluating the outcomes of any teaching program, so too are there two general ways in which a guidance program may be evaluated. First, a survey may be made of the procedures used in that program

with the purpose of determining the probability that the program is achieving the goals it is supposed to achieve. This method has been the usual one for evaluating guidance programs and goes back to a proposal made by Myers (13) in 1926. This is the traditional way of evaluating an educational program and one which is rapidly becoming outmoded in most educational fields outside of guidance. It is unsatisfactory as a technique principally because it is valid only insofar as definite knowledge exists concerning the extent to which specific procedures achieve specific goals. It assumes that much knowledge has been accumulated concerning the validity of guidance procedures, but since that assumption cannot be accepted, the survey method of evaluating guidance must also be rejected. It should be noted that the main reason why the survey technique for evaluating an educational program has been largely rejected in most fields of education is that it has been shown again and again that, at the present time, it is impossible to make valid guesses of what the outcomes of a program actually are. It has been found too often that the accomplishments of educational programs are much less than teachers commonly assume them to be. It is of vital importance to distinguish between evidence of the achievement of objectives and hopes that objectives are being achieved.

The second method of evaluating the outcomes of an educational program arises very largely out of a belief that the consequences of educational practices cannot be determined adequately unless evidence of those consequences is systematically collected. In this second method the procedure is that of defining carefully the objectives that are to be achieved, specifying the group in whom they are to be achieved, developing instruments for measuring the extent to which these objectives are achieved, and finally carrying through the program and then measuring its actual outcomes. One of the most striking changes in education in the last thirty years has been the almost universal change from the first method given above to the second method. It is a change from a prescientific method, which is likely to be influenced by wishful thinking, to a method in which wishful thinking plays very little part and in which assumptions are reduced to a minimum.

In all of this change from the prescientific to the scientific

method of evaluating outcomes, the field of guidance has played little part. In this respect, guidance programs have lagged far behind the times and with few exceptions represent the traditional rather than the modern approach to education.

*Evaluating Outcomes in Terms of the Achievement of Objectives*

A major difficulty in obtaining from guidance workers lists of outcomes of guidance is that such workers commonly believe that there is only one possible set of goals towards which guidance can possibly be oriented and that all guidance workers are attempting to achieve the same goals. This is a basic fallacy, for in guidance as in other areas of education, numerous different goals are possible and many of these goals are mutually incompatible with each other. Guidance programs aspire to produce desirable citizens, but the concept of what constitutes a desirable citizen varies from one person to another. Just as the term "desirable citizen" may mean entirely different things to different people so, too, does the term "adjustment" have many different meanings. A person who is well adjusted from the point of view of a union leader may be looked upon as a person in need of psychotherapy from the point of view of a business executive. Adjustment is worthy enough as a goal of guidance but useless as a concept unless it is operationally defined in great detail. Consequently, the statement that the goal of counseling is to improve the adjustment of the individual is about as useful as stating that the purpose of education is to produce educated citizens. This latter fact is one of the major reasons why so little has been done to evaluate the outcomes of guidance and only a few writers seem to have discussed this basic matter. Among the few is Lafferty (9) who has pointed out that practically all school guidance programs now lack clearly defined objectives, that they overemphasize the sheer mechanics of counseling as an end in itself, that they rely too much on the use of objective test scores, and try to fill too many needs. Tyler (24) and Wrenn and Darley (28) have also pointed out that the crucial steps in the formulation of rational guidance programs still have to be taken, namely, the definition of objectives. These writers agree that until the objectives of guidance have been clearly defined that little can be done to evaluate outcomes.

The same kind of difficulties in other areas of education has made the development of evaluation studies a slow and laborious business. As a matter of fact, it was only during the 1930's that methods were evolved for defining educational objectives in terms which made evaluation practical. However, the development of these methods for defining objectives has formed the basis for numerous evaluation studies which have thrown light on what learnings occur in certain specified situations. When the same methodology is applied to the type of learning situations provided by guidance workers useful evaluation studies will emerge.

For these reasons, it is not possible to review evaluations of guidance by listing a series of well-defined objectives of guidance and then presenting the evidence showing the extent to which each objective is achieved by given procedures. There is nevertheless some value in examining some of the evaluative criteria that have been used by various investigators. While few of these studies provide evidence of the efficacy of guidance or counseling for producing specific kinds of learning, they are worth reviewing because they illustrate some of the difficulties involved in evaluating guidance procedures.

Evaluative criteria fall into two general categories, subjective and objective. While objective criteria are in general much more satisfactory than subjective criteria, the latter must be considered because of the frequency with which they are used.

#### *Evaluative Criteria*

Subjective evaluative criteria of the outcomes of guidance include the individual's own assessment of his personal happiness, the satisfaction which he derives from his job, the extent to which he feels that his social life is adequate, and the degree to which he feels that he has achieved the goals which he set for himself. The chief difficulty in measuring these factors is that adequate instruments have not yet been made for their measurement and that the responses to the usual type of rating scale are too frequently colored by immediate and transitory circumstances. Day-to-day variations in job satisfaction and general personal happiness are large, and a minor catastrophe may temporarily color a person's entire outlook on life. Consequently, ratings by the individual of his feelings of satisfac-

tion or dissatisfaction with his job or his home life must be considered as highly invalid measures of the outcomes of guidance unless they are made on several different occasions.

Other subjective evaluative criteria include the satisfaction which a student feels with the counseling. This criterion has been used very commonly for evaluating counseling procedures. Studies by Mellon (12), Compton (4), and Paterson and Clark (14) all showed that a large fraction of those counseled reported that they found the treatment helpful. However, it can hardly be conceded that feelings of satisfaction with counseling can be considered either a major goal of the procedure or evidence of its success. The mere fact that the counselee feels satisfied is not evidence of the desirability of the process. People tend to be remarkably well satisfied with fortune tellers and other charlatans and tend to feel that they have derived much from the association. On a similar basis, one must reject as evidence of the efficacy of counseling, statements by those counseled that they had benefited by the treatment.

An argument commonly used for the validity of counseling procedures is that several studies have shown that students who cooperate with counseling are more likely to be well adjusted on the follow-up than those who do not cooperate. This has been cited again and again as evidence for the validity of counseling procedures, but it cannot be accepted as evidence, for the mere fact that a person cooperates with a counselor is in itself evidence of the ability to make adjustments of a certain kind. Similarly, an uncooperative attitude or a negativistic attitude toward the counselor is evidence of an inability to adjust to relatively simple social situations. Insofar as adjustment is a general factor, these follow-up studies show that those who cooperate with the counselor are going to be better adjusted than those who do not, but this does not imply that the counseling procedure was good, bad, or indifferent.

In general, it seems that the unsatisfactory nature of subjective criteria for evaluating counseling makes it necessary to use objective criteria, but these too must be used with great caution.

Objective criteria for evaluating the outcomes of guidance have included academic grades, income after a certain number of years, frequency with which jobs are changed, the stability

of life goals, the extent to which educational plans are completed, and so forth. Some of these criteria will now be examined to illustrate the caution that should be exercised in their use.

The criterion that has probably been most commonly used for evaluating the outcomes of guidance at the college level is the change in the average grades received by the student before and after counseling. This criterion is often based on the wholly unwarranted assumption that a major goal of counseling is to permit the student to improve his grades. With this goal in mind the counselor is tempted to seek out for the failing student the easy courses and the lenient instructors. While this practice may often result in an improvement in the student's grades, it does not result in an improvement in his work.

It is quite obvious that counseling and guidance should not serve the purpose of steering the student through college by showing him all the weak points in the administrative regulations and in the assignment of grades. Such a system serves the purpose of obtaining degrees for students who have not achieved the outcomes which such degrees are supposed to indicate. The counseling procedure which aims at improving grades by steering the weak student through the administrative and educational loopholes makes a farce out of the educational process for it makes him and others feel that he has achieved something which he has not.

This discussion does not imply that improvement in grades as a result of counseling occurs only when the student is shown loopholes in the educational system. However, it does imply that certain misguided workers raise the grades of their counselees by methods which assume that good grades rather than desirable behavior are the goals of clients. This does not mean that educationally respectable methods of helping the student to achieve more in his work do not exist. Time-old recipes such as those of helping the student to plan a schedule, helping him to improve his reading skills, helping him to choose a program which calls upon his outstanding talents, may all be effective ways of enabling him to achieve more than he would otherwise achieve. Where the latter techniques are used, *one* consequence may be a change in the grades of the counselee.

This discussion serves to point out that changes in school

grades can be used for evaluating the outcomes of guidance only when the objectives of guidance have been properly defined and where the outcomes of guidance are integrated with the outcomes of other aspects of education. An improvement in grades as a result of counseling cannot be accepted, without additional data, as evidence that the counseling process achieved worthy ends.

Since many of the implied goals of guidance are long-term matters, the measurement of the extent to which these goals are achieved must also be carried out over a long period of time. For example, much of the work undertaken in counseling is related to the formation of life goals, and it is quite evident that considerable time is necessary in order to determine whether these life goals are appropriate. In these long-term goals, as in short-term goals, there are both subjective and objective evaluative criteria which must be considered though, in general, long-term studies enable the investigator to use objective criteria.

Test scores are useful only insofar as they can be used to predict behavior in some specific situation in the future. If such predictions can be made, then at least one aspect of the guidance process has validity. Studies of the value of test scores for making predictions over short periods are too numerous to be summarized here. Studies of the validity of test scores for making predictions over long periods are fewer in number but much more significant for guidance workers. The long term studies undertaken by Thorndike and Lorge (19, 17, 20, 21, 18) are the most comprehensive studies of the latter kind. These investigators sought to determine whether tests administered early in secondary school could be used for predicting various aspects of the student's subsequent career. These investigations concluded that although educational guidance seems both possible and fruitful, the correlations of test scores with vocational success were so low that little worthwhile vocational guidance could be undertaken on the basis of test scores alone. Various criticisms (15, 7, 8) were made of the Lorge and Thorndike studies, some to the effect that guidance or counseling should not involve predictions of subsequent success, but the fact seems to remain that if test scores are



used in the guidance process then they must be used for making predictions.

Terman's follow-up studies of gifted children also yield data on the extent to which certain kinds of predictions can be made from certain kinds of test scores. In general, the Terman results are rather more promising than the Lorge and Thorndike results but this is probably a consequence of the fact that the Terman study was based on a group of extreme deviates.

An objective method of appraising counseling which is promising but which has been little used is the method of determining the extent to which behavior becomes reoriented towards more attainable goals as a result of the counseling procedure. An example of this technique is given in a study by Abramson (1) who found that those who were not judged to be suited for semi-professional or managerial work but who planned to enter it would frequently modify their life goals after counseling, but that those who planned to enter the professions modified their vocational goals less easily. There is a real need for studies of this kind which are fairly easily undertaken. Much could be done to appraise both changes in the life goals of the individual and in the understanding which the individual has of his own abilities. Counselees could be asked to rate themselves on various characteristics both before and after counseling. If test scores are discussed during the counseling interview, it should be possible to determine whether this procedure develops understanding in the individual of his own strengths and limitations. However, measuring techniques such as these should be used not only immediately before and immediately after counseling but also after an interval has elapsed. The fact that insight is achieved through counseling does not mean that the insight is permanent.

A subjective variation of this latter technique has been developed by Rogers and his associates at the University of Chicago. This technique requires the counselor to arrive at a subjective judgment of whether the client has developed insight into his problem. These nondirective counseling advocates believe that observation of the client will determine the degree of adjustment that has been achieved. The criticism of this method is obvious.

*The Use of Control Groups*

One of the basic difficulties in evaluating the outcomes of guidance is in finding suitable control groups. Control groups are important in the measurement of educational outcomes because they help the investigator to identify the cause of a particular outcome. It should be noted that in order for the control group to serve its purpose, it must be similar in all important respects to the experimental group.

In the field of guidance, evaluation has been undertaken on many occasions by comparing the behavior of those who received guidance with the behavior of a control group which did not. Unfortunately, there are hardly any studies on record in which the control group and the group receiving guidance (experimental group) were adequately matched. The common tendency has been to match control and experimental groups on the basis of irrelevant factors. For example, one study was carried out (27) in which the investigators studied a group that received guidance at the University of Minnesota Testing Bureau and an allegedly matched group which received no special guidance because the members of the group did not apply for any. The matching in this study was based on factors which had little relation to the purposes of counseling. Since the groups were compared in terms of their later adjustment, the control group and the experimental group should have been matched initially in terms of adjustment, and since this was not the case the outcomes of the experiment become almost impossible to interpret. It is hardly surprising under these circumstances that the counseled group showed better adjustment than the noncounseled group for, by applying for counseling, they had shown that they were individuals actually seeking an improvement in their adjustment to life. In this experiment the only meaningful control group would have been another group of individuals who by their behavior showed that they were actually seeking to improve their adjustment and which were not given personal counseling services.

One of the few published studies appraising a counseling program through the use of an adequate control group is that

by Toven (22). In this study 376 freshman college students were divided into two groups in order of registration. One group was counseled systematically throughout a four-year general academic curriculum. The other group had the same curriculum but did not receive special counseling. Counseling was undertaken by the usual faculty advisers which makes the study particularly interesting since the professional counselor commonly assumes that the work of the academic adviser bears little fruit.

The results of this study are important. Of the counseled group, 53.7 per cent graduated, but of the noncounseled group, only 36.2 per cent graduated. The counseled group seems to have had fewer academic difficulties though the two groups had almost identical academic averages during the last three years of college. The main effect of the counseling procedure seems to have been that it enabled more of the group to finish college. Of course, it must be remembered that any group singled out for special treatment is likely to be better motivated than those who fall in the run of the mill. One cannot help wondering whether the outcomes of this experiment may not be at least partly attributable to that factor. However, if the counseling procedure succeeded only in making the student feel wanted and appreciated it achieved an important end.

Another common and fallacious method of selecting a control group is to compare the subsequent behavior of those counselees who followed advice with those who did not. Studies by Webster (26), Burt (2), Macrae (11), Williamson and Darley (5), McConn (10), Viteles (25), Earle (6), Seipp (16), Trabue and Dvorak (23), and Clark (3) all follow this procedure. In each of these studies the validity of the "advice" given was tested in terms of whether those who followed the "advice" did "better" in some way than those who did not follow it. These studies neglect the fact that the person who rejects the advice of a counselor may be exhibiting a basic personality problem which may interfere with his success regardless of the situation in which he may find himself. In these studies, as in the one previously discussed, the basic difficulty arises from the fact that the control and experimental studies were matched for irrelevant variables. In these latter studies the only basis of

matching lies in the fact that the two groups to be compared applied for or were given counseling.

It is quite evident that there are major difficulties in the way of selecting adequate control groups for validating counseling procedures. At the present time the only theoretically satisfactory method of selecting a control group is usually administratively impractical. That method is to provide counseling services to alternate cases. This procedure provides two groups in which the desire for better adjustment is to some extent equated and would certainly provide a much more adequate control group than that usually selected on the basis of rather irrelevant material.

#### *The Selective Publication of Evaluative Studies*

Another factor which complicates the interpretation of validation studies is the tendency for studies with negative results to remain unpublished. This is not a result of any deliberate policy to suppress information which is inconsistent with the investigator's own point of view but rather a tendency which Charles Darwin noted when he said that somehow he just happened to forget facts which were inconsistent with his main theory.

The tendency for only those studies which indicate positive results to be published has an interesting effect on the statistical results of those that are published. It has the effect of biasing statistical tests of significance in such a way that the statistical significance of differences is greatly overestimated. It is essential then to interpret published studies with this factor in mind.

Guidance is not the only field in which the selective publication of results biases statistical tests of the significance of differences. The same thing happens in all fields where workers are bound by strong emotional ties to certain outcomes. The same thing has happened in studies of traditional versus modern classroom procedures. Most of those who carry out such studies are vitally concerned with showing that the newer educational practices are better in some ways than the older practices. The result is that investigators show a remarkable absentmindedness about publishing those studies that produce

negative results. This is evidenced from the tendency for minor studies to show marked positive results but for the large studies to show very small differences between the groups studied.

### *Present Status and Outlook*

At the present time it is not unfair to say that the chief evidence of the effectiveness of guidance is the subjective evidence which the counselor accumulates as a result of his experience with clients. This evidence has only limited value since it is likely to be influenced by wishful thinking and other irrelevant factors. The difficulty of obtaining objective evidence of what is learned by the counselee has resulted in a paucity of objective evidence concerning either specific aspects of guidance and counseling or the process as a whole. The majority of studies fail to provide interpretable evidence largely because they fail to control one or more important variables.

Progress will be slow until guidance workers come to recognize guidance as a learning situation which can be investigated by the methods developed for investigating other learning situations. These methods involve the specification of the objectives of learning that are to be achieved, the specification of the means of achieving these objectives, the selection of criteria for determining whether the learning objectives have been achieved, and provision for the control of relevant variables. Until more studies of guidance are undertaken following these steps, there will be very little certain knowledge of what guidance is actually accomplishing.

### REFERENCES

1. Abramson, L. S. "Relation of Environmental Factors to the Level of Occupation Choice and to College Attendance: 49 cases of St. Paul, Minn., Jewish Welfare Association Vocational Service, Minneapolis, Minn." University of Minnesota, June, 1944 (Master's Thesis).
2. Burt, C. and others. *A Study in Vocational Guidance*. Industrial Fatigue Research Board, Report 33, London, 1926.
3. Clark, R. "An Experiment in Vocational Guidance," *Vocational Guidance Magazine*, I (1923), 93-94.
4. Compton, R. K. "Evaluation of Guidance through an Elementary Psychology Class," *Journal of Applied Psychology*, XXV (1941), 703-711.

5. Darley, John. *Student Personnel Work*. New York: McGraw Hill Book Co., 1937.
6. Earle, F. M. *Methods of Choosing A Career*. London: G. G. Harrop and Co., Ltd., 1931.
7. Kitson, H. D. "Getting Rid of a Piece of Educational Rubbish." *Teachers College Record* (Columbia University), XXXVI (1934), 30.
8. Kitson, H. D. "Vocational Guiding is not Fortune Telling." *Teachers College Record* (Columbia University), XXXV (1934), 372.
9. Lafferty, H. M. "Some Characteristics of Guidance." *American School Board Journal*, CIX (1944), 29-30.
10. McConn, Max. "Educational Guidance is now Possible." *Educational Record*, XIV (1933), 475-499.
11. Macrae, Angus. "A Follow-up of Vocationally Advised Cases." *Journal of National Institute of Industrial Psychology*, V (1931), 242-247.
12. Mellon, E. H. "Evaluation of a Group Guidance Program." *School Activities*, XIII (1942), 295-297.
13. Meyers, G. E. *Some Tentative Standards for Judging a Comprehensive Guidance Plan*. Pittsburgh: Department of Vocational Guidance, 1926.
14. Paterson, D. G. and Clark, K. E. "Students' Judgments of Counseling." *Journal of Higher Education*, XIV (1943), 140-142.
15. Paterson, D. G. and Macrae, A. "A Target for Critics: Thorndike's Findings Challenged." *Occupations*, XIII (1935), 19-29.
16. Seipp, E. "A Study of One Hundred Clients of the Adjustment Service." *Adjustment Service Series, Report XI*. New York: American Association for Adult Education, 1935.
17. Thorndike, E. L. "The Prediction of Vocational Success." *Personnel Journal*, XII (1933), 189-197.
18. Thorndike, E. L. "Rebounds from the Target"—More about "The Prediction of Vocational Success." *Occupations*, XIII (1935), 329-333.
19. Thorndike, E. L., Bregman, E. I., Lorge, I., Metcalfe, Z. F., Robinson, E. E. and Woodyard, E. *Prediction of Vocational Success*. New York: New York Commonwealth Fund, 1934.
20. Thorndike, E. L. and Lorge, I. "The Chimera of Vocational Guidance." *Teachers College Record*, XXXV (1934), 359-371.
21. Thorndike, E. L. and Lorge, I. "Retests After Ten Years." *Journal of Educational Psychology*, XXV (1934), 136-141.
22. Toven, J. R. "Appraising a Counseling Program at the College Level." *Occupations*, XXIII (1945), 459-466.
23. Trabue, M. R. and Dvorak, B. J. "A Study of the Needs of Adults for Further Training." *Bulletins of Employment Stabilization Institute*, Vol. 3, No. 3. Minneapolis: University of Minnesota Press, 1934.

24. Tyler, R. W. "Principles Involved in Evaluating Student Personnel Services." *Student Personnel Services in Colleges and Universities*. Chicago: University of Chicago Press, 1941.
25. Viteles, M. S. "Validating the Clinical Method in Vocational Guidance." *Psychological Clinic*, XVIII (1929), 69-77.
26. Webster, E. C. "A Follow-up on Vocational Guidance." *Journal of Applied Psychology*, XXVI (1942), 285-295.
27. Williamson, E. G. and Bordin, E. S. "Evaluating Counseling by Means of a Control-Group Experiment." *School and Society*, LII (1940), 434-440.
28. Wrenn, C. G. and Darley, J. G. "Pupil Personnel, Guidance, and Counseling." *Review of Educational Research*, XII (1942), 57-62.