

VALIDATION OF A SELECTION DEVICE FOR VOLUNTEER PROBATION OFFICERS

Thomas M. Kelley, Companion Counseling Program
Wayne County Juvenile Court & U. of Detroit

& Daniel B. Kennedy, University of Detroit

Introduction

One extremely important yet grossly neglected area of research on volunteerism relates to the development and validation of instruments useful in distinguishing and selecting potentially "successful" from potentially "unsuccessful" volunteers. The present study is concerned with the evaluation and validation of one such instrument recently developed at the Oakland County Juvenile Court in Pontiac, Michigan.

Early in 1971 the Oakland County Juvenile Court was awarded a federal grant from the Law Enforcement Assistance Administration to develop and expand its already existing Volunteer Case Aide Program. Basically, the program entailed the recruiting of community volunteers to work on a one-to-one counseling basis with delinquent and neglected Juvenile Court wards. The secondary goal of the Oakland County Volunteer Case Aide staff was to develop an instrument that would ultimately be capable of assessing and predicting the potential competency of the volunteer.

In carrying out this second objective, it was first hypothesized that the general competency of the volunteer could be assayed and predicted by analyzing the continuous decision-making process occurring in the volunteer-youngster relationship. An attempt was then made to replicate the decision-making process in a pencil and paper instrument by employing the critical incident technique developed by Flanagan (1954).

The critical incident technique does not provide solutions. On the contrary, it obtains a record of certain behavior. In the Oakland County study, the specific behavior analyzed was the volunteers' verbalizations of their counseling experience with adolescents.

More specifically, the Oakland County research staff brought small groups of volunteers together to discuss the relationship with their assigned adolescents. These focus panel sessions were loosely structured with volunteers being asked to discuss such topics as:

1) Situations or events in the relationship with their assigned child where things went extremely well; and

2) Situations or events in the relationship where things went extremely poorly.

Volunteers were asked to go into considerable detail on these and other similar topics, describing the exact situation, indicating what was said, what was done, how they felt about what was said and done, and what might have been said and done differently. Each of these sessions was taped and the volunteers' verbalizations of their counseling experiences were analyzed in detail. As a result of these small group sessions, an initial sample of 51 "items" was generated, developed and edited. Each item represented a relatively typical critical incident situation which mirrored a problem occurring in the client-volunteer relationship. The reader would then be asked to select one of four plausible alternative solutions to the problems presented. Below is an example of a typical item:

Your youngster, a 15-year-old girl (already on probation for truancy), is in a court detention facility as a result of being truant from home over the holidays. You have worked with her for several months previous to the truancy but she never initiated a contact. Today you receive a call from her. She says she has run away from the court facilities. She pauses, and then tells you that she just can't say what is on her mind. How do you respond?

a. You tell her you can't help her much if she continues such maladaptive behavior and urge her to turn herself in;

b. You ask her where she is and, if she will give you a chance, pick her up and talk with her about what she should do with her truancy situation;

c. You try to find out where she is and then tell the police to pick her up; or

d. You discuss with her over the phone the probable consequences of running away from the county facilities and the options now open to her. You also encourage her to think it out carefully and decide what is best for her.

The next problem which confronted the Oakland County project staff was the determination of a scoring procedure for the 51 items and a preliminary indication of the usefulness of the items. Because of criterion problems in relation to measurement of successful casework outcomes, an indirect measure of success was devised.¹ In the words of the Project Director, Dr. Richard Traitel (1972):

"A procedure had to be devised whereby we could determine what the 'correct' alternative to each situation was and if a sufficient number of the appropriate respondents to that situation agree on a 'correct' answer. We had here the problem of determining a preliminary criterion group for the instrument, and an obvious choice for a group would have been 'successful' volunteer; their answers in agreement would reflect the competency factor which we are interested in measuring. Unfortunately, there seems to be no acceptable direct way to measure 'successful;' no single objective criterion for this concept. Therefore, another tack was taken. It was assumed that an indirect reflection of this competency and success would be the variable of 'experience' as a volunteer. The more successful and competent volunteer is likely to be more experienced."

Therefore, the Oakland County staff randomly selected a criterion group of volunteers from a pool of active volunteers with at least six months experience working with youngsters in the program. This group was composed of sixteen males and twenty-five females, with a mean age of 37 years and a mean educational level of 15 years of school. The group was then administered the 51 items previously described. Analyses were then performed on the frequency of selection for each of the four alternatives per item. As a result of this analysis, 32 items were generated for which the criterion group agreed upon a particular alternative beyond the chance level. These 32 items were allotted to parallel forms of 16 items each; the forms being further matched on relevant situational variables.

The purpose of the present investigation is to broaden the scope of validation of the Critical Incident Response Test by collecting response data on the instrument through its administration to a group of student volunteers presently involved in a federally funded (L.E.A.A.) Companion Counseling Program at the Wayne County Juvenile Court in Detroit, Michigan. Similar to the Oakland County Volunteer Case-Aide Program, volunteers in the Wayne County project worked with delinquent youth on a one-to-one counseling basis. However, all volunteers in the Wayne County program were students receiving academic credit for their involvement from Wayne State University and Wayne County Community College. Using this group of college student volunteers, this study was concerned with the following areas of investigation:

1) To determine the validity of the Critical Incident Response Test for a sample of experienced student volunteers in the Companion Counseling Program at the Wayne County Juvenile Court;

2) To investigate the potential differences in response pattern for items of the Critical Incident Response Test between experienced and inexperienced student volunteers;

3) To determine the nature of the total scoring differences on the Critical Incident Response Test between experienced and inexperienced student volunteers; and

4) To determine whether or not separating the experienced volunteer group into black and white sub-groups had any effect on the mean scoring differences on the Critical Incident Response Test and to determine the nature of this potential effect.

Population

The total sample of student volunteers was made up of 82 undergraduate students from Wayne State University and Wayne County Community College. There were equal numbers of experienced (N=41) and inexperienced (N=41) volunteers. A description of the total sample of experienced and inexperienced volunteers on age, sex, race, school and experience is presented in Table 1.

The group of experienced student volunteers tended to be younger, have more years in school and have more black students and more females than the inexperienced student group.

A comparison of black and white experienced volunteers is presented in Table 2 for the variables of age, sex, school, education and experience. While both groups had an equal amount of counseling experience, the black student volunteers were somewhat older and had one less year of college. There was a slightly higher percentage of males in the white volunteer group.

Method

In January, 1972, both forms of the Critical Incident Response Test were administered to the inexperienced student volunteer group during the first class session of the winter semester. These students had been recruited for the Companion Counseling Program but had received no orientation or training prior to the administration of the CIRT.

The experienced student volunteer group was administered the CIRT in March, 1972, after completing four and a half months of involvement in classroom lectures, counseling and contact with their assigned adolescents.²

Procedure

The frequency of response for the four possible alternatives for each item of the CIRT was computed for the total experienced group and again for the total inexperienced volunteer group. Chi square tests were then

Table 4: Chi Square and Significance Levels for Critical Incidence Response Test Items (Inexperienced Group)

Item	a	b	c	d	N	χ^2	P
1	9	18	12	5	44	8.1	.05
*2	0	0	3	41	44	109.0	.01
*3	0	0	30	13	44	55.0	.01
*4	0	37	7	0	44	78.4	.01
*5	1	41	0	2	44	103.0	.01
*6	0	0	10	34	44	70.2	.01
*7	27	3	5	9	44	32.7	.01
*8	1	2	14	27	44	40.5	.01
*9	0	19	25	0	44	45.6	.01
*10	23	6	1	14	44	25.3	.01
*11	4	3	30	7	44	44.5	.01
*12	3	31	3	7	44	49.4	.01
*13	4	5	0	35	44	72.9	.01
*14	1	0	27	16	44	45.6	.01
*15	3	0	8	33	44	61.6	.01
*16	4	4	32	4	44	53.4	.01
*17	9	5	1	28	45	39.0	.01
*18	2	9	22	5	38	23.4	.01
*19	4	1	4	29	38	51.4	.01
20	11	4	18	5	38	12.6	.01
*21	25	5	2	6	38	33.0	.01
22	5	14	15	4	38	10.2	.05
*23	3	21	12	2	38	23.8	.01
*24	0	13	24	1	38	38.6	.01
*25	22	5	1	10	38	25.0	.01
*26	2	4	7	25	38	33.4	.01
*27	9	27	0	2	38	45.4	.01
28	6	14	5	13	38	6.6	NS
29	0	18	4	16	38	23.6	.01
*30	31	3	1	5	40	59.6	.01
*31	1	30	3	7	41	55.8	.01
*32	0	15	21	2	38	31.0	.01

*These item alternatives were accepted as significant on the basis of at least 50% of the respondents choosing the particular alternative and a chi square $p < .01$.

Table 3: Chi Squares and Significance Levels for Critical Incident Response Test Items (Experienced Group)**

Item	a	b	c	d	N	χ^2	P
1	12	17	10	5	44	6.7	NS
*2	0	4	11	29	44	44.9	.01
*3	1	0	30	13	44	53.2	.01
*4	0	43	0	1	44	119.0	.01
*5	0	38	0	6	44	90.5	.01
*6	0	0	10	34	44	70.2	.01
*7	22	8	2	12	44	19.3	.01
*8	0	0	15	29	44	52.9	.01
*9	1	5	38	0	44	89.5	.01
10	17	6	6	15	44	9.3	.05
*11	5	2	33	4	44	57.5	.01
*12	1	32	2	9	44	55.3	.01
*13	2	5	4	33	44	57.5	.01
*14	1	1	24	18	44	38.0	.01
*15	2	0	1	41	44	109.0	.01
*16	5	0	34	5	44	65.6	.01
*17	8	1	1	35	45	71.0	.01
*18	4	5	31	5	45	48.5	.01
*19	1	0	3	46	50	121.0	.01
*20	7	8	27	4	46	27.8	.01
*21	24	1	13	7	45	28.2	.01
*22	1	29	10	5	45	43.8	.01
*23	0	30	3	8	41	51.6	.01
*24	1	18	26	1	46	46.9	.01
*25	31	7	0	7	45	51.6	.01
*26	2	6	7	30	45	45.6	.01
*27	12	25	0	8	45	29.7	.01
*28	2	18	7	14	41	17.0	.01
*29	1	29	4	12	46	46.3	.01
*30	37	4	1	2	44	82.4	.01
*31	0	28	1	16	45	46.9	.01
*32	0	17	27	1	45	49.5	.01

*These item alternatives were accepted as significant on the basis of at least 50% of the respondents choosing the particular alternative and a chi square $p < .01$.

**Ns vary from item to item because subjects were drawn from a larger group of fifty volunteers. The total sample of experienced volunteers (N=41) described in Table 1 are those who answered all test items with one choice for each item as instructed. Some students failed to answer all items, but did answer most of them. Therefore, some subjects were included in this item analysis, but not in the analysis of total mean scores.

computed for both groups on the frequency of selection of each of the four possible alternatives. The purpose of this computation was to determine those items which were answered in a non-chance fashion (i.e., other than equal distribution of choices across the four alternatives).

Two requirements had to be met before an item was judged significant: a significance level $<.01$ and a frequency of selection of the particular alternative greater than 50 percent.

Next, total scores for both forms A and B were computed for the experienced and inexperienced volunteer groups. Then, mean scores and standard deviations were computed for experienced and inexperienced sub-groups. T-tests were computed on the mean scoring differences between the experienced and inexperienced groups. This was done to determine the ability of the total scores on the Critical Incidence Response Test (rather than individual items) to distinguish between experienced and inexperienced student volunteers.

Finally, the total experience group was separated on the variable of race. T-tests were computed on the mean scoring differences between the racial sub-groups. This comparison was made as an attempt to determine the possibility of variables other than "experience" affecting the potential between group scoring differences on the Critical Incidence Response Test.

Table 1: Comparison of Experienced and Inexperienced Student Volunteers

	Experienced Group (N=41)	Inexperienced Group (N=41)
Average Age	26	29
Race: Black	18	36
White	23	5
Sex: Male	26	21
Female	15	20
School: WSU	24	0
WCCC	17	41
Education (mean years)	14	13
Experience (months)	4.5	0

RESULTS

Validity of the Critical Incident Response Test for the Wayne County Sample

Table 3 presents a breakdown on the frequency of selected item alternatives for the experienced student volunteer sample. The results indicated that for 30 of the 32 items, the experienced volunteer group agreed on a particular response alternative at a beyond chance level. Only two of the items (Item 1 and Item 10) did not reach significance. Furthermore, for all of the 30 significant items, the most frequently chosen alternative was the same as that chosen by the Oakland County development sample. Thus, for the Wayne County student volunteer group the Critical Incident Response Test displayed a high level of concurrent validity.

Table 2: Comparison of Black and White Experienced Student Volunteers

	Black (N=17)	White (N=23)
Average Age	32.8	25.4
Sex: Male	9	17
Female	8	6
School: WSU	5	19
WCCC	12	4
Education (mean years)	13	14
Experience (months)	4.5	4.5

Comparison of Item Response Patters for Experienced and Inexperienced Student Volunteers

Table 4 presents a breakdown of the frequency of response alternatives for the inexperienced student volunteer group. The results for the inexperienced group indicated a significant level of agreement on individual response alternative for 26 of the 32 items. Items 20, 22, 28 and 29, while reaching significance for the experienced group, did not show a significantly agreed upon alternative for the inexperienced group.³ Item 10, conversely, was significant for the inexperienced group but not for the experienced group.⁴ Item 1 did not reach significance for either group. The remaining 26 items were significant for both the experienced and inexperienced groups for the same response alternatives in all cases.

Thus, while there was somewhat greater variation of item response choices for the inexperienced group, the overall differences did not result in a powerful level of discrimination between the experienced and the inexperienced volunteers.

Comparison of Mean Scoring Differences Between The Experienced and Inexperienced Groups

Table 5 presents a t-test computed on the total mean scoring differences on the CIRT between the experienced and inexperienced volunteer groups. The results indicated that the experienced group scored significantly higher on the CIRT than the inexperienced volunteer group. Thus, while the individual items of the CIRT had little discriminative power between groups, the experienced volunteer

Table 5: Mean Scoring Differences Between Experienced and Inexperienced Volunteers

Group	Mean	SD	t
Experienced (N=41)	22.1	3.58	2.92*
Inexperienced (N=41)	19.5	4.35	

* p = .01

group, when compared to the inexperienced group, did display significantly higher total mean scores on the CIRT.

Mean Scoring Differences Between Black and White Experienced Volunteers

Table 6 presents a t-test computed on the CIRT mean scoring differences between white and black experienced student volunteers. The results revealed that the white experienced volunteer sub-group scored significantly higher on the CIRT than the black experienced sub-group. This finding would seem to suggest the possibility that factors other than "experience as a volunteer" might have influenced total scores on the CIRT, since the white and black students had exactly the same amount of experience.

As an indirect check on this possibility, a t-test was computed on the mean scoring difference on the CIRT between the experienced and the inexperienced black volunteer groups (see Table 7). The results of this comparison revealed no significant mean scoring difference on the CIRT between the experienced and inexperienced blacks. It should be pointed out that these sub-groups were quite similar on the background variables of age and education, and with the exception of five experienced blacks, all came from the same community college.

This finding would seem to lend support to the hypothesis that variables other than just "experience as a volunteer" were important in determining the higher total mean scores for the experienced volunteer group. It will be recalled, in this context, that the black experienced volunteers tended to be older and have less education than the white experienced volunteer group. It is possible that the greatest amount of education for the white volunteers bolstered their scores on the test, resulting in the significance between group scoring differences.⁵

Furthermore, it could be hypothesized that the age variable might be a factor affecting test scores since the younger white experienced student volunteers might have been more capable of identifying with situations involving adolescents like those presented in these test items.⁶ Or, it is possible that the black student group might have approached certain items on the CIRT with unique problem solving sets. These approaches, while possibly quite adequate, might have led to item choices different than those considered correct by the almost entirely White Oakland County development sample. In conclusion, whatever

Table 6: Mean Scoring Differences for the Experienced Group by Race

Group	Mean	SD	t
Black (N=18)	20.4	3.77	2.34*
White (N=23)	23.6	2.43	

*p = <.01

extraneous variables may have affected test performance, the assumption that "experience as a volunteer" was the primary cause of the significantly higher test scores for the experienced group would seem questionable to say the least.

CONCLUSIONS AND IMPLICATIONS

From the results of the present investigation, the following conclusions can be drawn:

1) The critical incident response test displayed a high level of concurrent validity when administered to the experienced Wayne County student volunteer group. For 30 of the 32 items, the experienced Wayne County test group agreed beyond chance on the same response alternatives as the Oakland County development sample.

2) On the whole, individual items on the CIRT were not powerful discriminators between experienced and inexperienced volunteers in this study.

3) Significant total scoring difference on the CIRT were found between the experienced and inexperienced student volunteer groups. However, further scoring comparisons between black and white sub-groups of the experienced group, and experienced and inexperienced black volunteers, revealed that the significant mean scoring differences between the two groups were probably influenced to some extent by variables other than just "experience as a volunteer." The results indicated that other variables such as age, and more probably, amount of education may be important in affecting performance on the CIRT for particular sub-groups.

The primary goal of the volunteer case-aide staff in Pontiac, Michigan, was to develop an instrument capable of distinguishing between

Table 7: Mean Scoring Differences for Experienced and Inexperienced Black Volunteers

Group	Mean	SD	t
Experienced Blacks (N=18)	20.4	3.77	.87
Inexperienced B Blacks (N=36)	19.4	3.97	

potentially successful and unsuccessful volunteers working in a one-to-one counseling type relationship with delinquent or neglected juveniles. In their efforts to develop such an instrument, the assumption was initially made that the more experienced volunteers would also tend to be more successful. The results of the present investigation would certainly question the validity of this assumption. Irrespective of that issue, however, it would seem clear that the next step that must be

taken is to develop and collect more objective criteria of volunteer success other than length of "experience" alone. Criterion measures that might be considered are: supervisory ratings of volunteers, ratings of volunteers by adolescents, pre- and post-comparisons of adolescent attitudes and personality measures, school grades, attendance, or even recidivism rates. Using such criteria, a true predictive validity study of the Critical Incidence Response Test must be undertaken. Such a study is presently under way at the Wayne County Juvenile Court in Detroit, Michigan, utilizing a completely new sample of student volunteers. The results of this study will be forthcoming.

Traitel, Richard

1972 "Preliminary Report on Development of the Critical Incident Response Test."
Unpublished, submitted to L.E.A.A.

Footnotes:

¹The chief test of many delinquency prevention and rehavilitation programs is what happens to delinquency rates. This is a poor test for two reasons. On the one hand, delinquency rates are an undependable index of the amount of delinquent conduct in a community. They go up or down with changes in law and with changes in community attitudes toward children's conduct, etc., as well as with changes in the actual amount of delinquent behavior. On the other hand, insofar as the rates are dependable, they register the joint effects of many factors in addition to those with which a particular delinquency prevention program is concerned. Control over these factors is difficult to achieve (Kelley and Kennedy, 1972 p.28).

²There was no sample mortality. All who began working with their assigned delinquents were still doing so when the test was administered. Each student contacted his case at least once per week for a minimum of three hours.

³It should be noted that each of these items reached significance at the .01 level but was not selected at the required 50% level.

⁴The response alternative chosen most frequently was different than the alternative reaching significance for the Oakland County sample.

⁵It should be remembered that the inexperienced group was composed almost entirely of black students (see Table 1).

⁶However, it could also be argued that older students would be more capable due to higher maturity and experience levels.

References

- Flanagan, J.C.
1954 "The Critical Incident Technique."
Psychological Bulletin, Vol.51, p.327-58.
- Kelley, Thomas M. and Daniel B. Kennedy
1972 "Delinquency Prevention and College Student Para-Professionals: A New Approach for our Juvenile Courts."
Volunteer Administration, Vol VI, 2,
p.28.