8-1978

Comparison of Work Samples and a Psychological Test in a Vocational Evaluation Program

Charles Dircken

COMPARISON OF WORK SAMPLES AND A PSYCHOLOGICAL TEST IN A VOCATIONAL EVALUATION PROGRAM

by

Charles Dircken

A Thesis
Submitted to the
Faculty of The Graduate College
in partial fulfillment
of the
Degree of Master of Arts

Western Michigan University
Kalamazoo, Michigan
August 1978

ACKNOWLEDGEMENTS

In writing this thesis, I have benefited from the encouragement, advice, and constructive criticism of Professors Frederick Gault, Chris Koronakos, and William Richardson. My thanks go to them, as to the many others at Western Michigan University, who have given much needed help. Appreciation is also extended to Leonard W. Lee, Director of State Technical Institute and Rehabilitation Center for permitting this research to be conducted at this facility, making subjects available for testing, and his constructive suggestions. Responsibility for the accuracy of all data obtained and reported rests entirely with me.

Charles Dircken

TABLE OF CONTENTS

INTRODUCTION

Work samples are evaluation tools utilized to assess an individual's skills and vocational potential. These tools are used extensively in vocational evaluation programs serving persons with various physical and emotional impairments so as to qualify for Federal-State Department of Vocational Rehabilitation Services. "The primary impetus for work sample evaluation derived from growing awareness that psychological testing was an ineffective method for evaluation of a large minority of the disabled population and that an alternative technique was needed." (Pruitt, 1970)

Although psychological testing has been criticized as effective assessment techniques, many are still widely used and advocated (Sankovsky, 1975; Botterbusch, 1973;1976). Gordon (1967) determined there was no significant differences in the validity of predictions favoring theory-oriented clinical assessment over work sample measurements, psychometric measurement, or an abbreviated assessment approach. What appears lacking in the criticism of psychological tests is a clear definition of what is a psychological test.

A survey of 15 facilities in Michigan indicated all but one used various psychological tests, aptitude tests, achievement tests, and personality and adjustment tests. The rationale for not utilizing psychological tests at the one state operated facility appears to be based on philosophical theory without any endeavor to substantiate or evaluate the work sample technique.

The present research compared predictions made from client participation on work samples, the Crawford Small Parts Dexterity Test (CSPDT), and a modification of standard administration methods for the CSPDT. Work samples selected for comparison were Dental Technology, Jewelry Repair, Office Machine Repair, and Watch and Clock Repair since all of these require fine finger dexterity and eye-hand coordination in working with small parts. These are skills purportedly measured by the CSPDT. Significant discrepancy might be anticipated in predictions between the three testing methods if one procedure were superior to the other(s).

METHOD

Experimental subjects were drawn from two successive enrollment groups in a vocational evaluation program to which they were referred by the State of Michigan Department of Vocational Rehabilitation Services. Persons who, in the regular evaluation process, participated in the Dental Technology, Jewelry Repair, Office Machine Repair, or Watch and Clock Repair Work Samples automatically were tested with the CSPDT and the Crawford Small Parts Dexterity Test-Modified (CSPDT-Modified). Subjects participating in two or more work samples were tested only one time on the CSPDT and CSPDT-Modified. This resulted in 37 subjects, aged 18 to 56 (median age 26) who participated in 49 work sample administrations serving as the experimental group.

Eighteen students enrolled in the Dental Technology, Office Machine Services, and Watch and Jewelry Repair training programs served as the control group. Six persons were randomly selected from each of these training programs. Because of time constraints imposed by instructional staff, only the CSPDT-Modified was administered to the control group. This was satisfactory for control purposes since the purpose of this research was to determine if there were significant discrepancies between predictions made from three testing procedures.

Work samples were designed by facility staff to predict training potential for the respective training and have been utilized for three or more years. Administration methods and norms have not been formally written and rely essentially on the individual administrator's experience. Despite this lack of standardization, confidence in these testing

tools is high with most staff members at this facility and they remain, esentially, the only assessment tools utilized.

The CSPDT is a standardized performance test designed to measure fine eye-hand coordination. It is comprised of a board containing 42 holes on the left side, 42 holes on the right side, and three bins for pins, collars, and screws across the top portion. For the one part, the examiner uses tweezers to pick up one pin, place it in a hole, and to pick up one collar to fit over the pin. This procedure is done six times for the top row on the right side of the board for practice. The process is then repeated for the remaining 36 holes on the right side of the board while the examiner times the examinee with a stop watch. Screws are inserted, one at a time, in the holes on the left side of the board and screwed all the way down using a screw driver. After completing the six holes at the top for practice the examinee is again timed while repeating this procedure for the remaining 36 holes. Norms have been established for various sample groups and are published in the 1956 Revised Manual by the Psychological Corporation.

For CSPDT-Modified the timed portions of the CSPDT were repeated four more times. Norms used for both the CSPDT and CSPDT-Modified were male trade and technical high school students. When repetition required more than one and a half hours testing was terminated, even if five trials were not completed.

RESULTS

Of the 49 work sample administrations there were 11 recommendations for potential to satisfactory complete specified training programs. Had the CSPDT-1 Trial (Standard Method) administration only been employed, only one of those persons would have been rated as having the necessary manipulative skills. Utilizing the CSPDT-Modified Procedure resulted in five of the eleven subjects being recommended and one as being questionable. Two other subjects would have been recommended using CSPDT-Modified and five rated questionable that were recommended by the work samples. Figure 1 lists predictions by work samples for all subjects and provides comparisons to the prognosis offered by the two methods of administering the CSPDT.

There were 20 discrepancies between results from the work samples and those of the CSPDT, using either method. These discrepancies are shown in Figure 2.

Since Watch and Clock Repair and Jewelry Repair Training constitutes one training department, there were only three training departments from which to draw control subjects for the CSPDT. Comparison of instructors' ratings and predictions from the two CSPDT administration methods are shown in Figure 3. Agreement between instructors' ratings and CSPDT-1 Trial was very low while CSPDT-Modified extremely high.

One year after obtaining these results it was discovered that only one of the experimental subjects had entered any of the three training

Figure 1: Comparison of predictions from work samples, CSPDT-1 Trial, CSPDT-Modified.

FIGURE 1

COMPARISON OF PREDICTIONS FROM WORK SAMPLES, CSPDT-1 TRIAL, CSPDT-MODIFIED

Experimental Group

Training Programs															
Dental Tech	Subject No.	4	7	8	9	11	12	14	16	22	24	27	30	35	37
	Work Sample	Yes	?	No	Yes	No	Yes	No	No	No	No	Yes	No	No	No
	CSPDT-1 Trial	No	?	No	No	No	No	No	No	?	No	?	No	Yes	No
	CSPDT-Modified	No	?	No	Yes	No	No	No	No	?	No	Yes	No	Yes	No
Jewelry Repair	Subject No.	4	6	7	9	11	13	15	20	22	25	27	32	33	
	Work Sample	No	Yes	?	Yes	No	No	No	No	?	No	No	No	No	
	CSPDT-1 Trial	No	Yes	?	No	No	No	No	No	?	No	?	No	No	
	CSPDT-Modified	No	Yes	?	Yes	No	No	No	No	?	No	Yes	No	No	
Office Machine Service	Subject No.	2	3	5	10	23	28	29	31	34	36				
	Work Sample	No	No	No	No	No	Yes	No	No	No	No				
	CSPDT-1 Trial	No	?	?	No	No	No	No	No	No	No				
	CSPDT-Modified	No	?	No	No	No	No	No	No	No	No				
Watch & Clock Repair	Subject No.	1	2	5	17	18	19	21	22	26	27	28	29		
	Work Sample	Yes	No	No	Yes	No	No	No	No	?	Yes	Yes	No		
	CSPDT-1 Trial	No	No	?	No	?	No	No	?	No	?	No	No		
	CSPDT-Modified	?	No	No	No	?	?	No	?	?	Yes	No	No		

Figure 2: Subjects where a discrepancy existed between predictions from work samples, CSPDT-1 Trial, and CSPDT-Modified.

FIGURE 2

SUBJECTS WHERE A DISCREPANCY EXISTED BETWEEN PREDICTIONS FROM WORK SAMPLES, CSPDT-1 TRIAL, AND CSPDT-MODIFIED

Training Programs										
	Subject No.	4	9	12	22	27	35			
	Work Sample	Yes	Yes	Yes	No	Yes	No			
Dental Tech	CSPDT-1 Trial	No	No	No	?	?	Yes			
	CSPDT-Modified	No	Yes	No	?	Yes	Yes			
	Subject No.	9	27							
Jewelry	Work Sample	Yes	No							
Repair	CSPDT-1 Trial	No	?							
	CSPDT-Modified	Yes	Yes							
	Subject No.	3	5	28						
Office Machine	Work Sample	No	No	Yes						
Service	CSPDT-1 Trial	?	?	No						
	CSPDT-Modified	?	No	No						
	Subject No.	1	5	17	18	19	22	26	27	28
Watch &	Work Sample	Yes	No	Yes	No	No	No	?	Yes	Yes
Clock Repair	CSPDT-1 Trial	No	?	No	?	No	?	No	?	No
	CSPDT-Modified	?	No	No	?	?	?	?	Yes	No

Figure 3: Comparison of ratings by instructors, CSPDT-1 Trial, and CSPDT-Modified.

FIGURE 3

COMPARISON OF RATINGS BY INSTRUCTORS, CSPDT-1 TRIAL, AND CSPDT-MODIFIED

TRAINING DEPARTMENT STUDENTS

Training Programs							
	Subject No.	D1	D2	D3	D4	D5	D6
	Inst. Rating	?	No	No	Yes	Yes	?
Dental Tech	CSPDT-1 Trial	?	No	No	Yes	No	No
	CSPDT-Modified	?	No	No	Yes	Yes	Yes
	Subject No.	W1	W2	W3	W4	W5	W6
Jewelry &	Inst. Rating	Yes	Yes	Yes	?	?	Yes
Watch & Clock	CSPDT-1 Trial	No	?	?	No	No	No
Repair	CSPDT-Modified	Yes	Yes	Yes	?	?	Yes
	Subject No.	OM1	OM2	OM3	OM4	OM5	OM6
Office Mach-	Inst. Rating	Yes	Yes	Yes	?	Yes	Yes
ine Service	CSPDT-1 Trial	?	?	No	No	Yes	Yes
	CSPDT-Modified	Yes	?	Yes	Yes	Yes	Yes

programs. This was Subject 4 who entered Dental Technology Training based upon recommendations derived from the Dental Technology Work Sample. This subject did not complete the training program as was predicted from both the CSPDT-1 Trial and CSPDT-Modified.

Looking at the status of the 18 individuals from the three training departments one year later indicates the CSPDT-Modified accurately predicted 12 yes' and no's. For five questionable individuals we found four either completed the programs or were still participating satisfactorily. The remaining questionable rated student transferred from Dental Technology to another training program within the facility where finger dexterity and eye-hand coordination are also indicated as being important. Listings of each control subject are shown in Figure 4 by training department.

Had we assumed that questionable ratings on the CSPDT-Modified should be given the benefit of the doubt, then predictions would have been correct in 88.8% of the control group. This would also have meant that nine more experimental subjects would have been given the opportunity to enter the training programs if CSPDT-Modified were criteria for recommendation rather than work samples.

Of the 33 experimental subjects who completed the CSPDT-Modified, five obtained a higher rating as compared to CSPDT-1 Trial ratings, while one subject received a lower rating as a result of the CSPDT-Modified. By contrast, 11 of the 18 training department students achieved a higher rating on the CSPDT-Modified than they achieved on the CSPDT-1 Trial.

Figure 4: Status of 18 training department students one year after testing.

FIGURE 4

STATUS OF 18 TRAINING DEPARTMENT STUDENTS ONE YEAR AFTER TESTING

Training Programs							
	Subject No.	D1	D2	D3	D4	D5	D6
	CSPDT-Modified	?	No	No	Yes	Yes	Yes
Dental Tech	Completed or Still Enrolled in Training Prog.	No*	No	No	Yes	Yes	No
	Subject No.	W1	W2	W3	W4	W5	W6
Jewelry & Watch & Clock Repair	CSPDT-Modified	Yes	Yes	Yes	?	?	Yes
	Completed or Still Enrolled in Training Prog.	Yes	Yes	Yes	Yes	Yes	Yes
	Subject No.	OM1	OM2	OM3	OM4	OM5	OM6
Office Machine Service	CSPDT-Modified	Yes	?	Yes	?	Yes	Yes
	Completed or Still Enrolled in Training Prog.	Yes	Yes	Yes	Yes	Yes	Yes

* Transferred to another training program within the facility.

Individual ratings for the CSPDT for the 37 experimental subjects are listed in Figure 5.

Figure 5: Ratings of 37 experimental subjects on the CSPDT.

FIGURE 5

RATINGS OF 37 EXPERIMENTAL SUBJECTS ON THE CSPDT

Subject Number	Percentile Ratings for Each Trial of the CSPDT										Predictions Based on CSPDT-1 Trial	Predictions Based on CSPDT-Modified
	Pins and Collars Trial					Screws Trial						
	1	2	3	4	5	1	2	3	4	5		
1	25	30	50	25	60	30	70	75	75	75	No	?
2	5	5	10	20	25	5	5	25	40	30	No	No
3	20	50	30	10	10	80	80	80	80	99	?	?
4	1	10	1	25	1	5	5	1	1	5	No	No
5	5	20	30	50	50	75	90	70	60	50	?	No
6	80	70	95	90	90	99	99	99	99	99	Yes	Yes
7	25	50	25	70	5	75	90	80	90	90	?	?
8	10	25	10	30	10	5	5	5	30	5	No	No
9	70	90	60	90	95	50	75	80	90	70	No	Yes
10	5	20	20	40	50	5	25	30	50	50	No	No
11	1	10	1	30	60	1	5	25	5	20	No	No
12	25	30	40	40	30	5	10	5	10	10	No	No
13	10	20	1	5	--	1	1	1	--	--	No	No*
14	5	5	5	5	10	1	1	1	5	5	No	No
15	5	10	10	40	--	1	1	1	1	--	No	No*
16	<1	--	--	--	--	1	--	--	--	--	No	*
17	10	20	30	30	40	50	70	70	75	60	No	No
18	25	10	30	30	40	80	75	75	90	90	?	?
19	5	10	20	10	10	5	75	95	95	75	No	?

FIGURE 5
(Continued)

RATINGS OF 37 EXPERIMENTAL
SUBJECT ON THE CSPDT

Subject Number	Percentile Ratings for Each Trial of the CSPDT										Predictions Based on CSPDT-1 Trial	Predictions Based on CSPDT-Modified
	Pins and Collars Trial					Screws Trial						
	1	2	3	4	5	1	2	3	4	5		
20	5	30	20	25	30	50	25	40	90	75	No	No
21	25	25	30	75	20	5	25	30	40	40	No	No
22	60	60	70	50	70	80	95	80	80	95	?	?
23	1	1	1	1	1	1	5	10	10	30	No	No
24	10	40	40	30	50	5	10	10	10	5	No	No
25	10	30	30	25	10	5	10	20	5	5	No	No
26	40	40	50	60	30	50	80	75	80	90	No	?
27	70	80	80	90	75	80	95	99	99	99	?	Yes
28	30	60	60	75	80	40	70	50	30	40	No	No
29	25	50	40	60	60	50	70	70	75	80	No	No
30	50	30	70	10	60	5	5	5	5	30	No	No
31	<1	1	1	5	5	1	5	5	5	25	No	No
32	<1	<1	--	--	--	<1	<1	--	--	--	No	*
33	<1	5	5	5	5	5	5	5	10	5	No	No
34	5	5	10	20	5	1	5	25	75	40	No	No
35	80	80	50	95	95	95	99	99	99	99	Yes	Yes
36	30	50	30	5	20	1	5	5	5	5	No	No
37	1	5	5	5	<1	<1	1	5	<1	1	No	No

DISCUSSION

Comparing ratings obtained from the CSPDT-Modified and the CSPDT-1 Trial appears to demonstrate that the modified technique provides more reliable skill measures for eye-hand coordination and finger dexterity for working with small parts. These results must be handled with caution since a small control group (n=18) was used and since there is no substantiation the work samples are reliable predictors of success for the training programs under consideration. Predictions made from performance on the CSPDT-Modified were in high agreement with instructors' ratings (88.8%) and continued success one year later for the control group. In view of the present findings, it does appear the CSPDT-Modified is a suitable method for determining skill measures from which to select or reject clients for training in these training departments and that those who are rated questionable by this testing procedure should be accepted for training.

Agreement between the four work samples and the CSPDT-Modified was lower (71.4%) but still significant enough to justify using it as a screening tool. Some of the evaluation clients were rejected for training because they lacked adequate reading skills, something not measurable with the CSPDT. By the same token, none of the work samples gave clear definitive measures of eye-hand coordination or fine finger dexterity. Less than 35% of the work sample results write-ups made mention of eye-hand coordination or finger dexterity and where these were mentioned, the statements were too general to be utilized for

making a prognosis for any other training or job placement.

Mean time for these work sample administrations was 30 hours per subject or a total of 1,470 hours for the 49 trials. The maximum time for CSPDT-Modified administration was 1½ hours or a maximum of 55.5 hours to test all experimental subjects. Since only 11 recommendations for training were made from the 49 work sample administrations, it must be concluded that these testing instruments still function to screen people out of entry into training programs. A screening-out process could have been accomplished in far less time.

The CSPDT was designed to measure an individual's current functional level and should be used cautiously as a prognosis for future ability. It is possible that some individuals do poorly on this type of testing because of not having engaged in activities where these skills were developed. Because of isolating a particular factor it could be possible to have additional information to aid in a habilitation program, information that does not appear to be generated from these work samples.

Work samples are utilized exclusively at the one State of Michigan-operated facility even though all other evaluation units utilize psychological testing and the literature by Sankovsky (1975) and Botterbusch (1973; 1976) recommends certain psychological testing. Standardized tests offer objective ratings so as to avoid the subjective ratings utilized on these work samples. Hamner et al. (1974) demonstrated that the sex and race of raters and ratees does affect ratings on simulated work sampling tasks. Observation and discussion with clients participating on these work samples also indicated a great deal

of anxiety, a criticism of work samples noted by Olshansky (1975). In some instances, anxiety was also noted on CSPDT administrations but this was usually alleviated after one complete trial--an additional benefit of using the CSPDT-Modified method.

Many tasks on these work samples required the subjects to learn skills rather than measuring entry-level skills. As was pointed out by Mushinski (1975) any knowledge, skills, and abilities that are to be learned during on-the-job training should not be included in the work sample. It is highly possible that including this type of learning on these tasks may inadvertently have resulted in measuring the evaluator's teaching skills rather than the subject's entry-level skills.

The problem could be resolved by designing work samples that minimize learning requirements and emphasize present functional skills' measures. Where instructions have to be provided they should be provided in a standard manner by all evaluators to all clients; utilization of audio-visual aids could assure standardized instructions. Deviations in providing standard instructions should be clearly itemized, as is required for all other standardized testing procedures. Utilization of standardized factorial tests, such as the CSPDT, should further facilitate assessing entry-level skills rather than the evaluator's teaching skills.

CONCLUSIONS

Based on data obtained during this project, the desirability of utilizing work samples to screen applicants for four of the training programs at one State of Michigan rehabilitation facility is considered questionable. An extensive amount of client time was utilized to participate on the work samples with only one person returning for any of these particular programs. This subject did not satisfactorily complete training. Since the work sample predicted success, while a modification of a standard psychological testing instrument indicated to the contrary, the reliability and validity of the work samples become more questionable. Because n=1 it is not possible to arrive at any conclusions as to the degree of reliability or validity of the work samples.

A control group of n=18 did demonstrate 88.8% agreement between predictions made for the CSPDT-Modified and student performance one year after testing. Similarly, comparisons of predictions made from one psychological test and the work samples were in high agreement. Since the CSPDT-Modified demonstrated higher prediction value than the CSPDT-1 Trial method, it does appear that the modified technique should be utilized for screening purposes rather than the administration method.

To obtain valid testing techniques to screen applicants for the training programs, it appears necessary to do extensive factorial testing of all students entering these programs and comparing these results to successful students and those who do not succeed in order to determine appropriate testing instruments to be employed in the future.

For reliability and validity purposes, it also appears crucial that objective rating instruments be used. Since this type of screening would take less time, much of the anxiety exhibited during work sample testing could be reduced since only persons expected to have a reasonable chance of success would participate on work samples which were matched to the client's skills and interests. Factorial testing also would have the advantage of providing specific information to be utilized in prescribing remediation necessary where an individual has interests but lacks skill(s).

When work samples are used, designers and administrators of these instruments must bear the burden of proof that the tests truly are reliable and valid. The current work samples did not meet these criteria nor did they have objective ratings. Failure to meet these qualifications can result in discrimination by the raters utilizing these tests which the results in further violations of court rulings and legislative actions to safeguard applicants and to promote effective rehabilitation programs.

REFERENCES

Botterbusch, Karl F., The Use of Psychological Tests with Individuals Who are Severely Disabled, Menomenie, Wisconsin, 1976.

Botterbusch, Karl F., Tests and Measurements for Vocational Evaluators, Menomenie, Wisconsin, 1973.

Gordon, Leonard V., Clinical, psychometric, and work-sample approaches in the prediction of success in peace corp training. Journal of Applied Psychology, 1967, 51, (2), 111-119.

Hamner, W. Clay, Kim, J. S., Baird, Lloyd and Bigoness, William J., Race and sex as determinants of ratings by potential employers in a simulated work-sampling task. Journal of Applied Psychology, 1974, 59, (6), 705-711.

Muchinski, Paul M., Utility of work samples. Personnel Journal, 1975, 54, (4), 218-222.

Olshansky, Simon, Work samples: another view. Rehabilitation Literature, 1975, 36, (2), 48-49.

Pruitt, Walter A., Basic assumptions underlying work sample theory. Journal of Rehabilitation, 1970, 36, (1), 24-26.

Sankovsky, Raymond, Evaluating Rehabilitation Potentials of the Severely Handicapped: Vocationally Related Components, Institute, West Virginia: Research and Training Center, 1975.