1977

Prediction of Industrial Rate in Handicapped Workers Using Subject Variables, Psychometric Data, Evaluator Ratings, and Work Samples

Joseph Schreiner
Loyola University Chicago

PREDICTION OF INDUSTRIAL RATE IN HANDICAPPED WORKERS

USING SUBJECT VARIABLES, PSYCHOMETRIC DATA,

EVALUATOR RATINGS, AND WORK SAMPLES

by

Joseph Schreiner

A Thesis Submitted to the Faculty of the Graduate School

of Loyola University of Chicago in Partial Fulfillment

of the Requirements for the Degree of

Master of Arts

September

1977

ACKNOWLEDGMENTS

I wish to thank Drs. Mayzner, Posavac, and Slaymaker for their suggestions and other help given during the course of this study. I also wish to acknowledge the services rendered for the computational work by the Loyola Data Processing Center.

TABLE OF CONTENTS

LIST OF TABLES

INTRODUCTION

One of the major tasks of a vocational evaluator is to predict the future vocational performance of mentally handicapped and other exceptional clients. To do this, the evaluator employs a battery of tests, structured work situations, and systematic observations. Depending on the criteria and the client population, a collection of predictor variables will have varying effectiveness in assisting the evaluator.

Several studies indicate a general ability factor that predicts subsequent work performance (Elkin, 1968; Levine & Elzey, 1960; Townsend, Prien, & Johnson, 1974; Wagner & Hawver, 1965). In such a situation, work performance can best be estimated by measuring the unitary general ability factor, with relatively little emphasis placed on other contributing factors. Generally, this is true if the client population is low-functioning. With low intelligence, a client cannot compensate for poor ability through other personal characteristics. A minimal level of ability seems to be necessary for the acquisition of basic work skills. Wagner et al. (1965) found that a general "intactness" factor pervaded all the predictor variables and criterion. Correlations among the variables were high. The subjects all had IQ's less than 50.

Conversely, in higher-functioning populations, general ability seems to be one of several components in predicting work performance

(Domino & McGarty, 1972; Levine et al., 1960; Sali & Amir, 1971; Shipe, 1971). One of the more important components is the personality of the subjects. Shipe (1971), in studying mildly retarded and borderline subjects (50 to 85 IQ), found that internality of locus of control and delay of gratification were correlated with vocational success for vocational school students. Domino et al. (1972), using four factors, found that work adjustment and personal adjustment were significantly correlated in a population of 35 young, female mental retardates. Their IQ's ranged from 58 to 79. In Israel, Amir et al. (1971) reported opposite results. In their study of low-functioning subjects (IQ's from 30 to 65), the author reported that personality characteristics were better predictors of vocational performance than IQ. However, their criteria consisted of two, three, or four discrete categories based entirely on ratings. The personality variables were also ratings made by the same personnel, confounding the results by a possible halo effect. The results were further confounded by the high correlations between the personality variables and the visual-motor tests (.58 to .73). IQ, itself, had respectable correlations with criteria (.24, .45, and .46).

Whenever continuous variables are used in a study, e.g., IQ, industrial rate, time to complete a work sample, a general ability factor usually emerges from the results (Elkin, 1968; Townsend et al., 1974). In contrast, when discrete variables are used, e.g., four-step ratings, success or failure at competitive employment, the results are hard to generalize and sometimes contradictory. This is true because continuous variables allow a more accurate assessment than discrete, all things being equal, and that discrete criteria from different studies may not be comparable. For example, Jackson (1973), in predicting eventual placement of mentally retarded adolescents into institutions, discovered that IQ was a valid predictor. Performance IQ superiority also indicated a better success rate. Fiester & Giambra (1972), however, found that verbal psycholinguistic ability was indicative of vocational success in adults. Kolstoe (1960) compared success and failure groups in competitive employment, and noted that appearance, lack of auditory handicap, good job skills, and cheerfulness were all predictive. Socioeconomic status, IQ, academic training, and urban background were not. McKerracher and Orritt (1972), in predicting outcome for a vocational training program, found sex and age related to success, but not IQ.

Other factors can predict work performance besides general ability and personality. One of these is visual-motor ability. Though closely related to intelligence in low-functioning populations, it emerges as a separate component in several studies (Sommers, Joiner, Holt, & Gross, 1970). Rosen, Kivitz, Clark, and Floor (1970) performed a factor analysis on both predictor and criterion variables and found a visual-motor factor among the predictor variables. The variables loading on this factor had significant correlations with the criterion variables. Levine et al. (1960) factor analyzed the San Francisco Vocational Competency Scale (SFVCS) and also found a visual-motor factor independent of general cognitive ability. Interestingly enough, the first factor extracted by principal axes was a general ability factor on which all variables loaded at least .40. Separate cognitive and visual-motor factors emerged after Varimax rotation.

Work habits and skills, such as punctuality, attendance, flexibility, persistence, and motivation form another important component in work performance. Unfortunately, work habits cannot be accurately assessed except through supervisor ratings. Nonetheless, these ratings often correlate well with work performance. Besides the cognitive and visual-motor factors, Levine et al. (1960) also found two factors related to work habits: flexibility and dependability. Using multiple regression, Song and Song (1967) predicted job efficiency in mental retardates. They found that variables measuring intelligence and work habits were the best predictors when used simultaneously. Bitter and Bolanovich (1970), in constructing the Work Adjustment Rating Form (WARF), found that their work skills ratings form correlated .60 with the objective criterion of work production.

An interesting theory integrating the roles of ability and work habits components in their contribution to work productivity was offered by Cohen and Close (1975). By experimentally manipulating the conditions of high and standard motivation, they found that the actual difference in productivity in standard motivation was related to the time spent not working, rather than differences in production rate. At high motivation, this effect was attenuated. This finding suggests that attending to task is the critical dimension in some workshop jobs, rather than ability. Differing production rates in equally able subjects may be accounted by their respective times actually working.

Interaction effects also play a role in work production. Brolin (1972) found that the adequacy of rehabilitation services could influence predictor variables. For those who had received adequate services as judged by three raters, almost all of the cognitive, visual-motor, demographic, personality, and work ratings were significant predictors of ultimate work performance. For those without adequate services, only age, performance IQ, and a few ratings were found to predict. They also found that males were more responsive to adequate rehabilitation services than females.

Besides the specter of different criteria, failure at cross-validation casts doubt upon the validity of some studies. Rosen et al. (1972), in twice trying to duplicate their previously mentioned study, were unable to obtain the same results, even with subjects from the same population. This effect was especially pronounced in the correlations between the predictor and criterion factors. The authors concluded that a shotgun approach for predicting work and community adjustment is not productive, and that personality measures, or variables with construct validity would be better predictors.

The adequacy of general cognitive ability tests, such as the Wechsler Adult Intelligence Scale (WAIS), Illinois Test of Psycholinguistic Abilities (ITPA), Stanford-Binet (SB), and Peabody Picture Vocabulary Test (PPVT) as predictors of vocational performance in mental retardates has been examined in several studies. Cochran (1970) standardized scores on the WAIS, SB, and PPVT so that they would represent equivalent estimates of mental ability in mental retardates. Norms for the WAIS were extrapolated to lower levels so as to be comparable to the other two. The PPVT, however, seems to be a weaker instrument than the WAIS or SB. Kaufman and Ivanoff (1968) found that the correlations between the PPVT IQ and the WAIS full scale IQ was only .17 within a workshop population. The Wide Range Achievement Test (WRAT) Reading Test, in contrast, correlated .50 with the WAIS IQ.

Subtests on the WAIS have also been used to predict work performance. Kaufman (1970) found the Comprehension Test to be the best discriminator between employed and unemployed mental retardates. The WAIS Arithmetic and WRAT Arithmetic tests were also good. Performance IQ superiority also seems related to community adjustment in mentally retarded adolescents, according to the same study.

The relation between academic achievement and mental retardation in workshops has been studied (Wallin, 1969). The author tested adult mental retardates in a workshop with the Wallin-Cutsforth Scale of Academic Achievement. He found wide variation in achievement. About one-third were illiterate, and about one-half were at least the third grade level. The average SB IQ was 53. Because of its easy administration, the WRAT is often given as a quick test in reading, arithmetic, and spelling. Its validity is good. Cochran and Petrini (1969) found that the WRAT subtests correlated well with the WAIS and SB, and moderately well with the PPVT. Atwell, Jamison and Fils (1969) administered the 1946 edition of the WRAT to 51 mentally retarded adolescents. One year later, they readminstered the test along with the 1965 edition. Correlation among the three administrations of each subject area were all above .91. This indicated reliability over both time and test forms.

Supervisor ratings are often used as predictor variables. With precautions, such ratings are sufficiently reliable, and, hopefully, valid measures. Abelson and Payne (1969) found that ward attendants could achieve good interrater reliability (above .80) on institutionalized mental retardates using objective items with two, three, or four alternatives. Lower reliabilities were attributed more to item, rather than rater, inadequacy. The poorer items consisted of adjectives, such as hyperactive, passive, or aggressive, instead of objective behaviors. Bitter et al. (1970), in developing the WARF, obtained an average reliability of .80 among four raters. The instrument, as mentioned before, correlated .60 with work performance in the shop.

A definitive statement as to the one best type of predictor of vocational success is nearly impossible. Generalizations among studies is difficult because of different criteria, predictors, and populations. However, a brief survey is possible. Gibson and Fields (1970), Rosen et al. (1970), Townsend et al. (1974), Song et al. (1967), and Kolstoe (1961) all examined a variety of predictors for vocational success. Their criteria ranged from sheltered workshop performance to placement in various success/failure groups, to job efficiency in competitive employment. Three criteria were dichotomous, the other two continuous. Each study examined at least three types of variables from the following five: cognitive ability, visual-motor ability, personality characteristics, work habits, and physical characteristics. Each of the five types had some success in predicting outcome. Cognitive ability was examined in all five studies; it was a successful predictor in three. Work habits and personality were both examined in four studies. Work habits were predictive in all four; personality in three. Visual-motor ability and physical characteristics were only examined in two studies. Visual-motor ability was significantly associated with success in both; physical characteristics in one.

An interesting finding indicating that the simple linear additive model may not be adequate for prediction was presented in Gibson et al. (1970). Here it was found that the one best predictor was the combined employment potential judgment of four rehabilitation professionals. However, the combination of IQ and social skills rating was equally effective. The authors hypothesized that the raters took both IQ and social skills and their possible interaction into account in making their prediction. By ignoring the interactive effect, previous studies may have underestimated the role of both variables in vocational success. A more efficient prediction system might be based on configurations or profiles, rather than simple addition or discrete client variables.

This thesis examines some of the variables mentioned previously: cognitive and visual-motor ability, physical characteristics, and work habits. Only one rating measures work habits, and there are no personality measures. On the other hand, sixteen structured work samples simulating actual jobs are used. As another category of predictors, they could prove enlightening.

Because this study is a secondary data analysis, there are no specific quantitative hypotheses to test. Much of the discussion will be devoted to a post hoc analysis of the results. However, there are several expectancies or trends that should arise from the proposed analyses.

First, as indicated by the literature, there should be a general ability factor expressed in the data. Variables likely to measure this are IQ, performance IQ (PIQ), dexterity tests, and an ability rating. Within the correlation matrix, there should be a positive manifold (all variables correlated positively). In a factor analysis, the first factor extracted through principal axes should represent ability and have a relatively large eigenvalue. The magnitude of this ability factor will be influenced by the reliabilities of the respective variables and the range of ability present in the subjects.

A work skills component of acquired work habits, relatively independent of ability, may also exist in the data. It should account for some of the variance in vocational performance that is not explained by ability. Its detection is dependent on the number and relative purity of variables measuring it. Unfortunately, there is only one "pure" measure of work habits in this battery, the evaluator's rating. The work samples should load moderately on this factor, though not as much as the rating. The standardized ability tests should not load significantly. Besides a factor analysis, partial correlation techniques may be able to elicit a work habits component by defining sources of variation. This can be done by examining the residual correlations between variables and criterion while controlling for the effects of the evaluator's ratings of ability and work habits.

The variables of this study can be categorized into four groups: subject variables, standardized tests (both cognitive and visual-motor), evaluator's ratings, and standardized work samples. Each of these groups have different characteristics and relations with criterion.

The subject variables include sex, age, and presence of a secondary handicap. In general, these should be poor predictors of work performance. The standardized tests, on the other hand, should be good predictors of criterion. These seven variables are likely to load heavily on any ability factor. Four of these variables measure cognitive ability, the other three visual-motor ability.

The evaluator's ratings should be excellent predictors of production, considering that they are global judgments based on many variables and systematic observation. However, the ratings cannot be expected to be perfect, because a subject is rated before he actually starts production on the work floor. Of theoretical importance will be the relation between the ability rating and the work habits rating. The ratings should be relatively simple variables in that they tap only one factor, i.e., the ability rating measures ability only, not fatigue, work skills, or previous experience. In contrast, the standarized work samples should be complex variables because of the many influences on their performance. In a factor analysis the work samples are likely to load significantly on more than one factor. However, like the ratings, the work samples should also be excellent predictors of work performance because they most closely approximate the criterion of actual production. Of vital interest will be the relative effectiveness of the ratings and the work samples in predicting performance. One type of variable represents the global judgment of the evaluator, the other a structured sample of relevant behavior.

From the perspective of general prediction, the optimal number of variables should be about three or four, although this is dependent on the number of factors involved in work performance. Should general ability be the only factor within the data, one or two variables might be optimal for prediction. Prediction should also be facilitated if the variables used are from different categories, e.g., a rating and a work sample should predict better than two work samples.

METHOD

Subjects

The subjects were 127 mentally retarded adults at Lombard Training Center of Chicago. They ranged in age from 16 to 45, with a mean of 23. IQ scores varied from 20 to 86, with a mean of 60. Forty percent of the clients were female. Subjects entered a four week evaluation program before being placed on the work floor. During this time, the vocational evaluator collected the data on each subject.

Variables

As previously mentioned, the variables in this study fall into four categories. Subject variables include sex, age, and the presence of a secondary handicap. Three dummy variables were constructed for the presence of a visual handicap, an emotional disturbance, and a physical handicap. Only 23% of the subjects had any secondary handicap.

The standardized test variables are the WAIS IQ and PIQ, the PPVT, a composite WRAT score to measure academic achievement, the assembly and composite scores from the Purdue Pegboard Test (PPT), and the total time for the Crawford Small Tools Dexterity Test. The rating variables are the vocational evaluator's ratings of ability, work habits, attainment of basic symbolic skills, and attainment of independent living skills. The final category of variables is the standardized work samples. A work sample is scored as the amount of time the subject needs to complete it. These tasks closely resemble the jobs that a subject is likely to encounter on the work floor. The sixteen tasks are: Flashlight Assembly, Lipstick Package, Name and Number Comparison, Screw Sort, Washer Sort, Color Pattern Collate, Washer Thread, Fiber Washer Sort, Aerosol Cap Sort, Pippette Assembly, Color Chip Sort, Color Discrimination Sort, Lid Inspection, Light Inspection, Packing Sample, and Slip Sample.

The criterion of vocational success for this study is the subject's industrial rate. Industrial rate is defined by the U.S. Dept. of Labor as the percentage of work output in goods and services a mentally retarded worker produces compared to the average worker in competitive employment. It is derived by dividing the subject's output by the amount a normal person could be expected to produce. Competitive production rates are obtained from either the company offering contract work to the sheltered workshop or computed from trial runs by rehabilitation workers in the shop.

Data Analysis

All data analysis was performed with the SPSS package on the IBM 370 computer at Loyola University of Chicago.

Several of the variables in this study do not have normal distributions. These are the work samples and the Crawford score. Their distributions are positively skewed to resemble those of reaction times. In order to render them linear normal and amenable to correlation coefficient analysis, a negative logarithm transform was performed on these scores.

Descriptive statistics were computed for all variables, including the mean, variance, standard deviation, skewness, kurtosis, and range. A correlation matrix for all variables was also derived.

Several factor analyses were performed on the variables. Both Varimax and oblique rotations were used. Canonical correlations were performed among the four groups also.

The more important multivariate analyses are multiple regression and partial correlation. Multiple regression with the criterion of industrial rate was performed on these groups of variables: the subject variables, IQ, PIQ, and academic achievement (AA); the three dexterity test scores from the Purdue and Crawford; the four evaluator ratings; the work samples; sex, age, PIQ, AA, and the dexterity tests; sex, age, PIQ, AA, and the work samples; the dexterity tests and the ratings; and the ratings and the work samples.

Partial correlations were computed for all variables with criterion while controlling for five other variables. The control variables used were PIQ, the Purdue Assembly score, the ability rating, the work habits rating, and the Screw Sort work sample. All possible first and second order coefficients were derived.

Missing Data

Because the data were not collected in a controlled setting, there is a considerable amount of missing data. No data among the subject variables or criterion is missing. Approximately 15% of the standarized test data and ratings is missing. About 50%, however, of the work sample data is missing. The total amount is sufficient to warrant a discussion of the reasons for the missing data and some methods for minimizing its effects.

There are two reasons for the considerable amount of missing data in the work samples. The first is random and does not bias the results. The repetoire of work samples administered changed slowly over the two years of data collection through deletion and addition of several work samples. As work samples wore out or lost parts, they were no longer administered. New samples would be created and administered in their place. Thus, clients who entered the program relatively late were not given several of the work samples that the earlier subjects were given, and visa versa. Missing data for the subjects in this case was dependent only on the time of entrance into the program, and independent of any other variables, such as sex, age, IQ, or handicap.

The second reason for missing data, however, does introduce bias. Lower-functioning subjects are unable to perform the more difficult work samples, and performed the others more slowly than the higher-functioning subjects. Consequently, the more difficult work samples were not completed, and so could not be scored. This resulted in a restricted range of subjects. These work samples should not correlate as well with criterion and other variables than the less restricted samples.

To partially quantify this bias, a new variable was generated, AV, the number of work samples the client did not perform. This variable should correlate negatively with the criterion and general ability. A multiple regression was also performed using AV and the work samples as predictors of industrial rate. First order partial correlations between all variables and criterion were also computed while controlling for AV.

Three options for handling missing data were considered: listwise deletion, pairwise deletion, and the substitution of a priori values for the missing data. The last alternative was rejected because of the differing reasons for the missing data. If low ability were the only reason for missing work sample data, substitution of arbitrary maximum values could prove adequate. However, work sample data is also missing for another, random reason, independent of ability. To substitute maximum values in these cases would add error to the work sample measurements, and thus obscure any relationships. Listwise deletion was rejected for another reason. Almost every subject has missing data in the work samples. In the more complex statistical analyses, there may be no subjects without missing data. To delete subjects with any missing data in that situation would be to delete the entire analysis. By default, pairwise deletion was used for the missing data option. The danger in this is the construction of a correlation matrix derived from different populations, especially the coefficients among the work samples. It should be noted that in SPSS, the degrees of freedom for the F ratio in multiple regression are based on listwise deletion, though the coefficients used may be based on pairwise.

RESULTS

Table 1 gives a summary of the descriptive statistics for several of the more important variables. Screw Sort and Color Collate are used as typical work samples because of their high correlations with criterion and their relatively large number of subjects. Note the differences in the skewness and the kurtosis between the original and transformed work sample data. These statistics indicate that the variables in this study are relatively normal and so appropriate for further analyses.

Table 2 illustrates some of the more important correlations with criterion. The subject variables are the poorest predictors, the work samples the best. Other good predictors included the dexterity tests, the ability rating, and performance IQ.

Table 3 shows a correlation matrix for nine of the most important predictor variables. As expected, there is a positive manifold (except for AV, which correlates negatively with all variables). Note that AV has near-zero correlations with the work samples. This implies a small amount of bias in the distribution of missing data through the work samples. AV also correlates negatively with ability measure, such as IQ, PIQ, and the ability rating.

Table 4 shows the first factor extracted through principal axes, the oblique factor matrix with three factors, and the estimated communalities. The factors all correlate positively with each other.

Table 1
Descriptive Statistics for Several Variables

Variable	Mean	Sta.Dev.	Range	Skewness	Kurtosis	No. of Cases
Industrial Rate	40.0	18.2	41.0	.88	.50	127
Age	22.7	5.4	31.0	1.66	3.41	127
IQ	59.5	16.3	71.0	.08	-.18	92
PIQ	59.8	17.2	69.0	.39	-.47	64
Academic Age (in years)	7.87	2.23	11.1	.49	.30	86
Purdue Composite	31.6	8.1	36.0	-.53	-.26	108
Purdue Assembly	20.9	8.2	36.0	.44	-.34	101
Ability Rating (0-9)	4.86	2.70	9.0	-.07	-1.03	108
Work Habits Rating	5.17	2.95	9.0	-.21	-1.27	111
AV(Missing Data from Work Samples)	10.0	2.95	11.0	.06	-.80	127
Screw Sort (Original)	83.8	46.7	269.0	2.08	5.88	82
Screw Sort (Transformed)	-1.87	.21	1.00	-.32	.10	82
Color Collate (Original)	76.3	36.0	189.0	1.18	1.32	77
Color Collate (Transformed)	-1.84	.20	1.00	-.07	-.15	77

Work Sample (original) units are minutes.

Table 2

Coefficients with Industrial Rate

	Coefficient	N
Sex (M=1, F=0)	.11	127
Age	-.10	127
IQ	.34	92
PIQ	.53	64
AA	.18	86
Purdue Assembly	.67	101
Purdue Composite	.58	108
Mentally Ill	.04	127
Visually Impaired	-.06	127
Physically Handicapped	-.06	127
Ability Rating	.63	108
Work Habits Rating	.39	111
AV	-.37	127
Screw Sort	.66	82
Color Collate	.60	77
Washer Sort	.50	33
Flashlight Assembly	.48	17
Pipettes	.56	50
Lipstick Package	.62	86
Washer Thread	.63	102
VQS (Verbal vs. Performance Ability)	-.40	64

Table 3

Intercorrelations Among Several Variables

1. IQ
2. PIQ
3. Academic Age
4. Purdue Assembly
5. AV
6. Ability Rating
7. Work Habits Rating
8. Screw Sort
9. Color Collate

Lower triangle: Correlation Coefficients
Upper triangle: N of Cases

	1	2	3	4	5	6	7	8	9
1		64	67	74	92	82	84	58	57
2	.94		46	55	64	58	58	42	38
3	.68	.63		70	86	81	82	59	54
4	.44	.49	.29		101	87	89	71	70
5	-.38	-.47	-.14	-.30		108	111	82	77
6	.54	.76	.24	.56	-.51		106	70	66
7	.06	.17	.09	.23	-.08	.26		71	68
8	.42	.68	.13	.48	-.12	.62	.25		63
9	.35	.60	.25	.48	-.13	.46	.31	.58	

Table 4

Factors from Oblique Factor Analysis

Variable	1st Factor Principal Axes Non-Rotated	FACTORS I	II	III	h^{2}*
IQ	65	46	92	25	90
PIQ	83	60	92	56	98
Purdue Composite	68	83	37	38	72
Purdue Assembly	72	84	48	38	74
Crawford Tools	69	72	50	38	56
Ability Rating	78	65	66	67	76
Work Habits Rating	33	32	02	33	17
AV	-36	-30	-56	-14	34
Lipstick Package	74	69	34	68	61
Screw Sort	83	69	40	76	74
Washer Sort	74	46	32	84	86
Color Collate	63	61	33	46	42
Washer Thread	77	80	48	48	66
Fiber Washer Sort	71	51	25	93	87
Aerosol Cap Sort	68	62	36	44	58
Pipettes	74	67	40	51	65
Industrial Rate	80	79	33	65	70

*Sum of squared factor loadings may exceed h^2 because of correlated factors.

Two subclusters of general ability resulted, a visual-motor and a cognitive factor. The other factor seems to load heavily with the work samples, especially those that involve sorting. In the original principal axes extraction, the first factor accounted for 49% of all variance; the second only 11%. This first factor was obviously a general ability factor, loading heavily on PIQ, industrial rate, and the screw sort. AV and the work habit rating were the two lowest loading variables. This latter finding lends support to the existence of an acquired work habits factor, or at least that this variable may be tapping something independent of ability.

Because of subject mortality several variables were deleted from all further analysis. They all had less than 40 subjects each. Their correlations were, in general, inconsistent with each other and often based on less than 10 subjects. The deleted variables were: the PPVT, Flashlight Assembly, Washer Sort, Color Chip Sort, Lid Inspection, Light Inspection, Packing Sample, Slip Sample, and Name and Number Comparison. All remaining coefficients were based on at least 30 subjects each, and so may be assumed to be indicative of the population parameters. All correlations with industrial rate were based on at least 45 subjects each.

Unfortunately, because of the missing data, the SPSS package was unable to perform canonical correlation among the several groups of variables. In every case the matrix was not positive definite. This unfortunate result makes interpretation of underlying factors within different variable categories difficult.

Table 5 summarizes the results of the various multiple regressions.

Table 5

Multiple Regression Summary

Variables Entered Hierarchically	R^2	F ratio (of entry)	df	F ratio (for 3 variables in equation)
1) PIQ	.24	14.13	1/44	16.57
AA	.29	8.90	2/43	2.42
Mental Illness	.30	6.03	3/42	.50
2) PIQ	.24	14.13	1/44	8.85
AA	.29	8.90	2/43	1.47
VQS	.30	5.90	3/42	.21
3) Purdue Assembly	.36	32.40	1/57	1.69
Purdue Composite	.40	18.32	2/56	2.95
Crawford Tools	.42	13.24	3/55	2.26
4) Ability Rating	.35	37.01	1/70	18.20
Work Habits Rating	.40	23.39	2/69	6.29
Living Skills Rating	.40	15.42	3/68	.08
5) Screw Sort	.48	6.58	1/7	.74
Washer Thread	.55	3.70	2/6	.71
Color Collate	.61	2.56	3/5	.68
6) Screw Sort	.48	6.58	1/7	1.67
AV	.56	3.75	2/6	.90
Lipstick Package	.62	2.68	3/5	.79
7) Purdue Assembly	.36	17.62	1/31	8.83
VQS	.42	10.65	2/30	1.48
PIQ	.44	7.65	3/29	1.38
8) Ability Rating	.35	21.15	1/40	4.66
Work Habits Rating	.40	13.22	2/39	3.90
PIQ	.41	8.85	3/38	.47
9) Purdue Assembly	.36	17.62	1/31	5.42
Ability Rating	.45	12.51	2/30	4.06
Work Habits Rating	.50	9.59	3/29	2.49
10) Screw Sort	.48	6.58	1/7	1.20
Washer Thread	.55	3.70	2/6	1.08
Work Habits Rating	.62	2.67	3/5	.83

In all cases, the first variable to enter the equation in hierarchical inclusion accounts for most of the predictable variance. Inclusion of other variables does not improve the regression equations much, though their entry F ratios are sometimes significant. An interesting development in the regressions was the necessity of deleting IQ and the substitution of VQS, or relative verbal over performance IQ strength. In all of the regressions where IQ and PIQ were allowed to enter the equation, the absolute values of their beta weights exceeded unity, a theoretical impossibility. Considering the IQ-PIQ correlation of .94, this abberant result was not surprising. In this sample, IQ and PIQ were practically equivalent measures. As such, they overdefined their relation to industrial rate and produced the deviant beta weights. To circumvent this, a new variable, VQS, was computed as IQ minus PIQ. This variable is the relative strength of the verbal IQ over the performance IQ. PIQ and VQS contain the same information as IQ and PIQ, but do not overdefine the variance in common with criterion when taken together. Therefore, they can be entered into the same regression without producing deviant beta weights. VQS itself correlates -.40 with industrial rate.

Table 6 shows the results of the first order partial correlations. The results were not definitive, but seem to indicate four pools or sources of variance. IQ, PIQ, AA, and the ability rating form a common source which can be labeled general or cognitive ability. The dexterity test scores from the Purdue and Crawford define another source, visual-motor ability. Work samples seem to have their own specific source of variance. Finally, the evaluator's rating of work habits seems to be

Table 6

Correlations with Criterion, Controlling for Several Variables

	CONTROLLING FOR					
VARIABLE	PIQ	Purdue Assembly	Ability Rating	Work Habits Rating	Screw Sort	AV
IQ	--	.08	.01	.33	.05	.22
PIQ	--	.28	.09	.47	.04	.40
AA	-.26	-.05	.00	.11	.06	.10
Purdue Assembly	.47	--	.41	.57	.43	.56
Crawford Tools	.39	.20	.30	.50	.21	.45
Ability Rating	.38	.38	--	.55	.28	.51
Work Habits Rating	.36	.32	.30	--	.31	.38
Living Skills Rating	.12	.10	.08	.33	.21	.30
Screw Sort	.57	.58	.52	.67	--	.70
Washer Thread	.48	.37	.46	.66	.36	.60
Color Collate	.45	.46	.47	.56	.36	.61
Pipettes	.49	.51	.47	.55	.36	.62

fairly independent of all other variables.

DISCUSSION

Before discussing the results of these analyses, it would be appropriate to evaluate the effects of the missing data and its bias.

AV, as the number of work samples not administered, is a rough quantitative measure of this bias. Its moderate coefficients demonstrate the pervasiveness of the non-random assignment of work samples and their limited ranges. Its most extreme correlation was with PIQ at -.47, followed by -.40 for the Crawford, and -.37 for the criterion. As expected, it had a negative relation with criterion and ability. The high relation with the Crawford is the result of the same subject mortality process. The Crawford is a manual dexterity test that requires the use of small tools. A portion of the subjects, about one-third, were unable to obtain a score on this test because they were unable to work with tools. Because AV was constructed to measure this dropout process in work samples, it is only natural that these two variables be correlated.

Though the effects of bias are apparent in the correlations with AV, the partial correlations reveal how small it actually is. The correlations of all variables with industrial rate are not affected by controlling for AV. Among the work samples, which should have shown the most effect of AV, there was no change in the criterion correlations. A few of the more important work samples had correlations lowered minimally. A few of the work samples had their correlations raised minimally. In both cases, the changes were so small that they could easily be attributed to statistical fluctuations. The work habits rating was completely unaffected.

When AV and the work samples were allowed to regress hierarchically onto industrial rate, the results also indicated little bias. Screw Sort entered the equation first with an R-square of .48. AV entered next, raising it to .55. At this point, the beta weights were .67 and -.27 respectively. The introduction of more work samples increased R-squared rather slowly, at .05 increments. The beta weights among the successively added work samples became more equivalent, levelling off at approximately .25. The interpretation of these results is clear; the work samples are all tapping the same variance, hence the low increments in R-squared and the equal beta weights. AV does not increase R-square any more than the addition of more work samples to the regression. As a measure of missing data bias, it does not help the predictability of criterion any more than adding relatively useless variables to the equation; its effects are minimal.

In contrast to the small effects of AV on the regressions and the partial correlations, the missing data devastated the canonical correlations. Of the six attempted analyses, only two were actually performed (those not involving the work sample variables). In the completed analyses, the significance levels of the first canonical variates were woefully close to chance expectation (alpha levels larger than .90 for each). These results can be attributed to the missing data and the pairwise deletion used in computing the correlation matrix. The four unprocessed analyses were rejected for lack of positive definiteness in the matrices. The lackluster results of the two completed analyses were caused by forcing some of the coefficients to zero to establish a suitable matrix for processing. Another bias expected from the missing data was the direct relationship between the strength of the work sample correlation with industrial rate and the number of subjects with scores on the work sample. This expectation was borne out. The four work samples with more than 75 subjects are the four highest correlations with criterion (all at .60 or above). The next three work samples with 50 to 75 subjects, have the next highest correlations (between .55 and .59). The remaining three, each with less than fifty, have correlations of .50 or below. The explanation for this phenomenon lies in the restricted range of subjects. The more difficult work samples are given only to the higher functioning subjects. This results in fewer subjects on these variables and less covariation of ability with criterion. Consequently, the smaller covariation is translated into smaller coefficients for the variables with fewer subjects. The trend, however, is not large. Eight of the ten work samples correlate with criterion from .50 to .65. The other two are anomolous, at .12 and .23.

The manipulations of AV and an inspection of the work sample correlations indicate the existence of a small amount of bias resulting from the distribution of missing data. AV is negatively related to industrial rate and ability, as expected. It has little effect on the other variables in this study, especially the work samples. The effects of restricted range on some work samples are also present, but apparently not large. The minimum number of subjects per correlation coefficient is sufficient to presume a correlation matrix approximating population parameters. The coefficients are all consistent with each other. The canonical correlations are the only casualties attributable to missing data. However, this seems to be more of a computational artifact brought on by pairwise deletion, than an inherent instability in the data.

A general ability component is present in the data. In the correlation matrix, all variables are positively related, except for the subject variables, and AV, which is negatively related to all. The various factor analyses all indicated an ability factor, loading with PIQ, Screw Sort, and industrial rate, accounting for approximately half of all variance.

The factor analyses and partial correlations decomposed this general ability factor into three smaller, correlated factors. The first is the visual-motor factor loading on the three dexterity tests and industrial rate. These three variables were shown to tap the same source of variance in the partial correlations. The relatively high industrial rate loading (.79) here seems to indicate that work production has a stronger relation to this component of ability. The second factor is a cognitive ability factor, loading on IQ and PIQ. IQ, PIQ, AA, and the ability rating were grouped together from the results of the partial correlations. Industrial rate loads lightly on this factor, indicating little relation to actual on-site production. The final factor is dominated by the work samples, mostly the sorting operations, with a strong loading from industrial rate. Of particular interest is the relative invariance of the ability rating across the three factors.

Apparently, all three components entered equally into the evaluator's global judgment of ability. As such, the rating is almost a perfect composite score of general ability across three related areas. Industrial rate, however, has unequal loadings across the areas, stronger in the visual-motor and work sample, and weaker in cognitive ability.

There is the possibility that these three factors are methods factors, rather than construct factors. As methods factors, their existence would be indicative of different measuring techniques, instead of actual hypothetical constructs. This possibility cannot be dismissed. However, as mentioned before, the ability rating loads well on all three factors and was computed through a completely different method. Because this variable is found in all three factors, it is safe to assume that they are not methods factors, and that they actually reflect different aspects of ability present in the data.

The existence of a work habits factor independent of ability cannot be inferred from the analyses. The one relatively pure measure of this construct is the work habits rating. It was predicted that this factor would have a heavy loading from the rating and moderate loadings from the work samples in a factor analysis. This was not the case. Though a work sample factor resulted from the analysis, there was no heavy loading from the rating. In fact, its loading was equivalent to the work habits loading on the visual-motor factor.

The work habits rating was relatively independent (uncorrelated) of other variables. Correlations with the work samples ranged from .06 to .45. Its correlation with the ability rating was .26; with industrial rate, .38. In the partial correlations, controlling for this rating had no effect on any of the correlations with criterion. Controlling for ability (the evaluator's rating) had no effect on the work habits rating-industrial rate correlation. When the four evaluator's ratings were allowed to regress hierarchically onto criterion, the ability rating entered first with an R-squared of .35. After the entrance of the work habits rating, the R-squared was only .40.

The work habits rating is related only minimally to ability measures. Though it correlates positively with most variables, it loads weakly on the three ability factors and the one massive ability factor. As such, it does not seem to be a "poorer" ability measure.

Given that the work habits rating is unrelated to industrial rate, ability, or any other variable, what can be said about it? There are three possible explanations for the results. The first is that the variable is unreliable or unstable. The trait it measures may change with time or place. The second explanation is that the rating has a large proportion of specific variance that no other variable taps. A third explanation, closely related to the second, is that there are no other good measurements of acquired work habits in the variables. Without these other variables, the underlying construct cannot be manifested in a factor analysis or a correlation matrix. Of these three explanations, the first is most likely to be true. If the work habits rating truly measured a factor independent of ability that contributes to work productivity, it would correlate well with industrial rate. It would also significantly raise R-squared in a regression equation when added to the ability rating. Neither of these situations exists. The work habits rating correlates only .38 with industrial rate and adds only .05 to the predictability of criterion when combined with the ability rating.

The subject variables provide little predictability for industrial rate, as seen from Table 2. Even the presence of a secondary handicap does not seem to influence production. There are three reasons for this. The first is the limited range of handicaps. All subjects entering Lombard Training Center have a primary handicap of mental retardation severe emough to warrant sheltered placement. Of all the subjects with a visual impairment, only one is actually blind. The others have a limited degree of vision. All of the emotionally disturbed subjects are at least stable enough to emit appropriate behavior on the work floor. Otherwise, they would not be in the workshop. In both of these variables, the degree of impairment is not enough to influence work performance beyond that of the primary handicap. The second reason is the difficulty in categorizing an impairment. Some, perhaps most, of the subjects had visual-motor difficulties, yet were not categorized as physically impaired. The official diagnosis for each subject were made by different physicians and psychologists, each with different definitions of impairment. Categorization was especially difficult for the presence of mental illness or a physical handicap associated with mental retardation. The final reason for poor correlation with criterion is the small number of subjects having a secondary handicap. The power of the correlation coefficient to measure a relationship was weakened by the unequal number of subjects either possessing or lacking a secondary handicap, e.g., 8 visually impaired subjects vs. 119 non-impaired.

The situation with the psychometric tests is better. The three dexterity tests correlate well with criterion and with each other, forming a tight cluster. They also correlate well with PIQ and the ability rating. The Crawford, especially, predicts well considering its restricted range and fewer subjects. The global academic achievement score, in contrast, predicts industrial rate poorly. By a simple path analysis, it can be seen that academic achievement is a predictor only to the extent it correlates with IQ, a moderately good predictor. The IQ-AA coefficient is .68; the IQ-industiral rate is .34. Multiplying these coefficients yields .22, the predictability of industrial rate by AA through IQ. The actual AA-industrial rate coefficient is .18, reasonably close to the theoretical .22. It appears that AA, in itself, has no relation to work productivity, except to the extent that both relate to intelligence.

The same phenomenon appears in the relation of IQ, PIQ, and industrial rate. However, a path analysis cannot be performed here because of a linear dependence between IQ and PIQ. PIQ, on the WAIS, is derived from the sum of the standard scores on five subtests. IQ is derived from the sum of these five tests and six others. This procedure raises the IQ-PIQ coefficient artifically. It can be seen, though, that PIQ is a distinctly better predictor of industrial rate than the more global IQ. VQS, or relative verbal over performance strength, correlates negatively with criterion, again indicating that it is the "performance" aspect of intelligence that influences work production, not the "verbal." IQ is a good predictor of industrial rate only to the extent that it taps performance, not verbal abilities.

The status of the two evaluator's ratings, ability and work habits, was discussed previously. The ability rating seems to be a stable and valid measurement tapping three correlated ability factors. It is a good predictor of industrial rate and has demonstrable relations to other variables. The work habits rating, though, is apparently an unstable variable not related to industrial rate, ability, or anything else. The other two ratings, symbolic skills acquisition and living skills acquisition, do not predict industrial rate any better than the ability rating. Their correlations with criterion disappeared when controlled for the ability rating. Correlations with other variables are moderate, and not particularly illuminating. The only exception to this is the correlation between the symbolic skills rating and academic achievement (.66). This is not surprising, since both are estimates of how well the subject can work with numbers and letters.

The work samples, though afflicted with the most missing data, seem to be the best predictors of industrial rate. In general, the work samples permitting the widest range of subjects to perform correlate best with criterion. They also correlate well with the other ability measures, PIQ and the dexterity tests. In the factor analyses, the work samples formed their own factor. The three highest loading variables on this factor are all sorting operations. These tasks involve visual discrimination of features and categorization of the objects. Though manual dexterity is important in the work samples, it is not as important in the work sample factor. Dexterity exists as another factor with heavy loadings from the Purdue and Crawford. Finally, it should be noted that the logarithmic transformation performed on the work sample scores was essential for this analysis. Without the transformations, the work sample correlations with criterion ranged from .14 to .41, instead of .50 to .66.

Theoretically, what can be summarized about this data? A massive general ability factor is found in the variables, accounting for 49% of all variance. It can be decomposed by factor analysis and partial correlations into three components: visual-motor ability, cognitive ability, and work sample/sorting ability. The evaluator's ability rating considers all three factors equally. However, the criterion, production rate in the workshop, correlates differentially with the three components. There does not seem to be a work habits component in these data. Finally, the best predictors of industrial rate seem to be the relatively simple structured work samples measuring the largest range of ability.

Practically, what do these analyses imply about predicting the industrial rate of mentally retarded clients in a sheltered workshop? Most important is that very few measures are needed to achieve an optimal prediction. In all of the multiple regressions performed in this study, the hierarchical entrance of a second variable into a regression equation never brought about an R-squared increment larger than .09. Most hovered at .05. Apparently, one score on an appropriate variable is almost as good a predictor as any linear combination of variables. This finding is a direct consequence of the general ability factor. The only good predictors of industrial rate are those that measure general ability, especially visual-motor. These good predictors all tap the same source of variance, and none of them is able to increase predictability by correlating with industrial rate through another source of variance. In other words, measuring general ability is practically measuring production rate.

Another suggestion from these analyses is the type of variables a vocational evaluator could use to predict work production. The superior variables were relatively simple structured work samples approximating typical jobs. Other good measures are dexterity tests, performance IQ, and the evaluator's rating of ability. Subject variables, such as age, sex, and secondary handicap, did not predict well.

Besides eliminating the missing data, several changes made in this study could clarify the relation between vocational success and other variables. Other criteria for success could be used. These might include daily attendance, ultimate competitive employment, or job satisfaction with workshop placement. Predictor variables, such as family background, personality, or need for money, could be used to inspect those aspects of work productivity not related to general ability. And finally, ratings from more than one evaluator could be contrasted against those from a single evaluator. A global judgment of two or more people may prove to be a more efficient way of predicting industrial rate than a series of work samples.

REFERENCES

Abelson, R., & Payne, D. Regional data collection in state institutions for the retarded: Reliability of attendant ratings. American Journal of Mental Deficiency, 1969, 73, 739-744.

Attwell, A., Jamison, C., & Fils, D. Relationship between the WRAT, a behavior guide, and achievement with retarded adolescents. American Journal of Mental Deficiency, 1969, 73, 879-882.

Bitter, J., & Bolanovich, D. WARF: A scale for measuring job-readiness behaviors. American Journal of Mental Deficiency, 1970, 74, 616-620.

Brolin, D. Value of rehabilitation services and correlates of vocational success with the mentally retarded. American Journal of Mental Deficiency, 1972, 76, 644-651.

Cochran, M., & Pedrini, D. The concurrent validity of the 1965 WRAT with adult regardates. American Journal of Mental Deficiency, 1969, 73, 654-656.

Cochran, M. A profile of psychological test scores for retarded adults. American Journal of Mental Deficiency, 1970, 74, 582-584.

Cohen, M., & Close, D. Retarded adults' discrete work performance in a sheltered workshop as a function of overall productivity and motivation. American Journal of Mental Deficiency, 1975, 79, 526-529.

Domino, G., & McGarty, M. Personal and work adjustment of young retarded women. American Journal of Mental Deficiency, 1972, 77, 314-321.

Elkin, L. Predicting performance of the mentally retarded on sheltered workshop and non-institutional jobs. American Journal of Mental Deficiency, 1968, 72, 533-538.

Fiester, A., & Giambra, L. Language indices of vocational success in MR adults. American Journal of Mental Deficiency, 1972, 77, 332-337.

Gibson, D., & Fields, D. Habilitation forecast in mental retardation: The configural search strategy. American Journal of Mental Deficiency, 1970, 74, 558-562.

Jackson, R. Prognostic significance of the performance-verbal ability patterns in predicting employment adjustment of EMR adolescents. American Journal of Mental Deficiency, 1973, 78, 331-333.

Kaufman, H. Diagnostic indices of employment with the mentally retarded. American Journal of Mental Deficiency, 1970, 74, 777-779.

Kaufman, H., & Ivanoff, J. Evaluating the mentally retarded with the peabody picture vocabulary test. American Journal of Mental Deficiency, 1968, 73, 396-398.

Kolstoe, O. The employment evaluation and training program. American Journal of Mental Deficiency, 1960, 65, 17-31.

Levine, S., & Elzey, F. Factor analysis of the san francisco vocational competency scale. American Journal of Mental Deficiency, 1960, 73, 509-513.

McKerracher, D., & Orritt, C. Prediction of vocational and social skill acquisition on a developmentally handicapped population: A pilot study. American Journal of Mental Deficiency, 1972, 76, 574-580.

Rosen, M., Kivitz, M., Clark, C., & Floor, L. Prediction of post-institutional adjustment of mentally retarded adults. American Journal of Mental Deficiency, 1970, 74, 726-733.

Rosen, M., Floor, L., & Baxter, D. Prediction of community adjustment: A failure at cross-validation. American Journal of Mental Deficiency, 1972, 77, 111-112.

Sali, J., & Amir, M. Personal factors influencing the retarded person's success at work: A report from Israel. American Journal of Mental Deficiency, 1971, 76, 42-47.

Shipe, D. Impulsivity and locus of control as predictors of achievement and adjustment in mildly retarded and borderline youth. American Journal of Mental Deficiency, 1971, 76, 12-22.

Sommers, P., Joiner, L., Holt, L., & Gross, J. Reaction time, agility equilibrium, and kinesioperceptual matching as predictors of intelligence. Perceptual and Motor Skills, 1970, 31, 460-462.

Song, A., & Song, R. Prediction of job efficiency of institutionalized retardates in the community. American Journal of Mental Deficiency, 1967, 73, 567-571.

Townsend, J., Prien, E., & Johnson, J. The use of the position analysis questionnaire in selecting correlates of job performance among mentally retarded workers. Journal of Vocational Behavior, 1974, 4, 181-192.

Wagner, E., & Hawver, D. Correlations between psychological tests and sheltered workshop performance in severely retarded adults. American Journal of Mental Deficiency, 1965, 69, 685-691.

Wallin, J. Educational status of clients in a workshop and training center for adolescent and adult mental retardates. The Journal of Genetic Psychology, 1969, 114, 41-62.

APPROVAL SHEET

The thesis submitted by Joseph Schreiner has been read and approved by the following Committee:

Dr. Mark Mayzner, Director
Professor, Psychology, Loyola

Dr. Frank Slaymaker
Assistant Professor, Psychology, Loyola

Dr. Emil Posavac
Associate Professor, Psychology, Loyola

The final copies have been examined by the director of the thesis and the signature which appears below verifies the fact that any necessary changes have been incorporated and that the thesis is now given final approval by the Committee with reference to content and form.

The thesis is therefore accepted in partial fulfillment of the requirements for the degree of Master of Arts in Experimental Psychology.

Oct. 3, 1977
Date

Mark L. Mayzner
Director's Signature