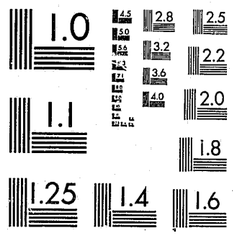

MF-1

SCREENING FOR RISK: AN ASSESSMENT OF THE ICFS PROJECT INSTRUMENTS

BY

FAYE S. TAXMAN

DON M. GOTTFREDSON

JAMES O. FINCKENAUER

June, 1981

s project was supported by Grant No. 78-NI-AX-0152
rded to the Research Center of the Graduate School
Criminal Justice, Rutgers, The State University,
ark, New Jersey. The funding was provided by the
tional Institute of Law Enforcement and Criminal
tice, Law Enforcement Assistance Administration,
Department of Justice. Points of view or
inions stated in this document are those of the
thors and do not necessarily represent the official
sition or policies of the U.S. Department of Justice.

This is one of a series of reports on the Improved Correctional Field Services Project Evaluation. The series consists of these parts:

1. Abstract

2. Executive Summary by Don M. Gottfredson, James O. Finckenauer, John J. Gibbs and Stephen D. Gottfredson.

3. The Improved Correctional Field Services Project: A Case Study by James O. Finckenauer and Don M. Gottfredson.

4. Screening for Risk: An Assessment of the ICFS Project Instruments by Faye S. Taxman, Don M. Gottfredson and James O. Finckenauer.

5. Risk, Supervision, and Recidivism: The First Six Months of Recorded Experience in the Improved Correctional Field Services Project by Don M. Gottfredson, James O. Finckenauer, and Faye S. Taxman.

   Appendix A: ICFS Instructions for Coding.
   Appendix B: Characteristics of the Sample for the First Six Months of Experience in the ICFS Project.

6. Social Adjustment: A Preliminary Report of the Improved Correctional Field Services Project by James O. Finckenauer and Faye S. Taxman.

7. The Needs and Concerns of Probationers: A Thematic Analysis of Interviews by John J. Gibbs.

8. The Needs and Concerns of Probationers: An Analysis of Questionnaires by John J. Gibbs.

9. Additivity and Interactions in Offense Seriousness Scales by Stephen D. Gottfredson, Kathy S. Young and William S. Laufer. 6572

10. Describing Probation Populations: Offense Seriousness by Stephen D. Gottfredson.

    Appendix A: Offense Seriousness Scoring System.

11. Exploring the Dimensions of Judged Offense Seriousness by Stephen D. Gottfredson.

    Appendix A: Offense Seriousness Study (survey form).
    Appendix B: The Question of Scale Value
    Appendix C: Replication of Factor Structures

SCREENING FOR RISK: AN ASSESSMENT OF THE ICFS PROJECT INSTRUMENTS

Contents

Tables

Figures

ABSTRACT

Three correctional field services (probation) projects-- in Kane County, Illinois, Suffolk County, New York, and Florida-- were developed and funded by the Law Enforcement Assistance Administration in 1978. The objective of the Improved Correctional Field Services Project (ICFS) was to determine the effectiveness of using risk screening procedures to assign probationers to different levels of probation supervision. Each agency was to employ a locally derived and validated risk screening instrument. The issue examined in this report is the validity of these risk screening devices.

Major methodological issues in the development and testing of risk screening devices are discussed. The instruments used by the participating agencies are examined in relation to these issues. Results of attempted validations of each, based on the first cohort sample of ICFS cases for the first six months of probation supervision are reported, and an exploratory study toward improvement in risk assessing is described.

Of these three sites, only one, Suffolk County, used an instrument with demonstrated substantial validity for Suffolk probationers. When applied to the first ICFS sample, results for each site indicate that validity must be seriously questioned. The limitations of sample size, short follow-

up, and possible confounding with treatment effects are such that further study of the ICFS risk screening procedures is much needed.

In the Improved Correctional Field Services project, each site employed a locally derived risk screening instrument.[1] These devices are key features of the project because they served as the main classification decision-making tools. A risk score was derived for each probationer, and this score defined a particular risk classification. In accordance with a quasi-experimental research design, staff at each site determined which persons were placed in each level of supervision.[2] The aim was to discern any differential impacts of varying levels of supervision for the different risk categories.

This report will consider the validity and utility of the instruments used by the agencies participating in the project. First, the nature of such instruments as used in probation settings is briefly reviewed.[3] Second, an overview of methodological issues and problems involved in constructing and validating risk assessment devices is presented. This discussion is intended to furnish a framework for the final section of the report--the examination and analysis of the three instruments, with an assessment of their validity. Finally, some exploratory work toward potential improvement of these risk screening measures is discussed.

THE NATURE OF RISK SCREENING DEVICES

Risk assessment instruments are decision-making tools that may have a variety of functions, including assistance in diagnosis, classification, and prediction of future behavior. These activities are based upon three key concepts that are "related but not identical."[4] An examination of these concepts and their application will help define the nature of risk assessment devices as applied in probation settings similar to the sites participating in the Improved Correctional Field Services project.

Diagnosis

Diagnosis, a term used originally in medicine, referred in that context to the problem of identifying a disease by means of its symptoms. Applied in probation settings, diagnosis would require the identification of some "signs" or "symptoms" which may be related to states of the probationers relevant to the aims of the probation program. The objective then would be to describe the probationer and to identify his or her needs. An example of a diagnostic tool is the Wisconsin Division of Correction's "Reassessment of Client Needs" form.[5] Using this device, the client is assessed in 12 areas: academic/vocational skills, employment, financial management, marital/family relationships, companions, emotional stability, alcohol use, other drug use, mental ability, health, sexual behavior, and the probation officer's impression of the client's needs. Each variable is considered to be a useful piece of information in identifying and assessing the needs of the probationer. The risk screening methods discussed in this report are not intended to be diagnostic tools in this sense.

Classification

Risk assessment devices can also serve as classification tools. Classification involves grouping like entities. The resulting categories may have either an empirical or theoretical basis, depending upon how the groups are derived and defined. Thus, a classification method may begin with a conceptual base, leading to the grouping of individuals into categories determined by psychiatric, sociological, or psychological theory. Alternatively, individuals may be grouped on the basis of certain empirically derived characteristics, the selection of which may or may not be theoretical.

Classifications can have many purposes, not all of which involve prediction. In probation settings, classification methods may be used to assign probationers to levels of supervision or types of supervision believed to be appropriate to their needs or for societal protection. The reasons for categorizing individuals into different supervision categories can include aims of efficient management of agency resources (for example, improved caseload management) or treatment purposes (for example, provision of specialized treatment pro-

grams for offenders with drug abuse histories). Thus, in view of the various uses of classification designs, classification instruments often reflect the aims of diagnosis or of prediction of future behavior. Diagnostic tools such as the Wisconsin case classification method may be used to make classification decisions. Similarly, prediction instruments such as the devices used by the participating ICFS sites may be used for classification purposes.

Prediction

Prediction is the primary aim of risk screening devices. Here the purpose is to forecast future behavior or future status of an offender or potential offender. In probation, the typical outcome of interest is success or failure on probation, usually defined in terms of arrest(s), conviction(s), and/or technical violation(s). Probationers are grouped (classified) on the basis of attributes known to differentiate "successes" and "failures," and to help predict the likelihood of these outcomes for similar groups.

It should be emphasized that when the classification purpose is screening for risk, as in the ICFS project design, the resulting classifications do not necessarily have any treatment relevance. That is, they do not necessarily provide diagnostic or needs assessment information that would provide guidance to treatment or supervision placement.[6]

METHODOLOGICAL ISSUES

Five general steps are usually followed in developing, testing, and using risk instruments.[7] First, criterion classifications (probation outcomes) are chosen. Second, a set of predictors or attributes of the probationers that will be used to predict the criterion are identified. Predictors in probation studies usually have included demographic items, characteristics of the criminal history of the probationer, and other probationer attributes. Next, a sample is selected that is believed to be representative of the probationer population to which the results will be generalized. The purpose of the sample is to assess the relation of possible predictor items to the criterion (and with one another) and to help select those items which appear to be potentially useful in prediction. The fourth step in the construction of an instrument is to determine how well it works. Verification or cross-validation requires the application of the instrument to new probation samples. Finally, if shown to be valid, the instrument can be applied with those probationers for which it was designed.

The history of prediction studies shows that these essential steps are often ignored or only haphazardly applied. As a result, risk assessment instruments tend to be plagued by problems stemming from flaws of method that can directly influence their utility. There are, in addition, a number of relevant concerns that instrument design-

ers should address. These include: (1) the "selection ratio problem"; (2) the "base rate problem"; (3) the fundamentally important issues of reliability and validity; (4) sampling concerns; and (5) the relative efficiency of various statistical methods used to combine predictions. These methodological issues will be discussed only cursorily here because there is an extensive literature on these subjects.[8] It is necessary to review them at least briefly, however, since our assessment of the ICFS project instruments refers to these issues as well as to the usual procedures for developing and testing such devices.

Selection Ratio Problem

The selection ratio refers to the number of people chosen to participate as a proportional sample of the total number available to participate in any given event. The usefulness of a risk assessment instrument for selection or assignment is related to this ratio. In instances where the selection ratio is low (that is, when only relatively few subjects are to be accepted into a program) an instrument with only low validity may be useful. If, however, a large proportion of the persons available are included in the sample, then the instrument must have greater validity in order to be useful.[9]

Base Rate

Another ratio influencing the potential usefulness of

a risk assessment device is the base rate. Base rate refers to the proportion of individuals in a target population who can be predicted to fall into a given category. It is the number of correct predictions that would be made if all members of a population (say, of probationers) were classified as expected to succeed. In probation studies, the base rate of success would be that proportion of probationers who succeed relative to the total number of probationers in the sample. Meehl and Rosen pointed out that the failure to consider base rates in developing instruments often makes it impossible to evaluate the usefulness of a prediction or risk assessment instrument.[10]

The developers of risk screening instruments should not be satisfied with their results if predictions can be made just as accurately without the instrument, from knowledge of the base rate alone. A prediction method may, however, be useful when it yields more information than that given by the base rate.[11]

With rare or infrequently occurring events, which lead to lower base rates, the difficulties of prediction are increased. This is due to the difficulty of finding predictors that will differentiate between categories in the criterion, because the variability in outcomes is small.

There may be other practical consequences of base rates. One is that two types of decision errors must be expected: (1)

some predicted successes will fail; and (2) some predicted failures will succeed. These errors, which involve social and monetary costs, are affected by the base rate. Further, devices are often used on populations other than those from which the original sample was drawn. Any problems from unwarranted generalizations to new samples may result in part from the presence of different base rates of success.

Ohlin and Duncan, who were among the first investigators to highlight the significance of the base rate problem, developed an "index of predictive efficiency" which assesses the relative utility of a prediction instrument.[12] Their index reports the percentage change in prediction errors from that provided by the base rate alone.

The index of predictive efficiency is related to another more recent statistic for comparing the relative efficiency of different prediction instruments -- the mean cost ratio (MCR). The MCR is a statistic that ranges in value from 0 to 1; the relative magnitude provides an indicator of successful prediction above that given just by the base rate. A procedure for significance testing has been developed on the basis of the relation between the MCR and a statistic called Kendall's tau.[13] This procedure increased its usefulness in determining whether or not the observed relation can be attributed reasonaly to chance. For these reasons, the MCR will be used in the assessment of the ICFS instruments.

Reliability and Validity

Issues of reliability and validity are fundamental concerns that are complicated by their relation to other methodological problems inherent in the construction of a risk assessment instrument: the representativeness of the sample, the classification of the criterion categories, the selection of predictors, the number of predictors included in the instrument, and the methods used in combining predictors. Reliability refers to attaining similar classifications over repeated observations or measurements (that is, the consistency of measurements), while validity refers to the degree to which the method achieves its predictive purpose.

The reliability issue must be addressed not only for the instrument as a whole but for each of its items; similarly, the reliability of the criterion classifications is an important concern. A useful instrument must discriminate among subjects, result in similar categorizations over repeated applications, and have demonstrably stable relations to the criterion categories.

The issue of reliability of the instrument hinges ultimately on the quality of the information used as predictors. It is axiomatic that no predictive device can be better than the information from which it is derived. Often the information on which predictors are based is not very reliable. Unreliability can arise from errors in the process of extracting data from probation case files or other records, or from errors in the original entries in the

case records.

The validity of an instrument is influenced by both the representativeness of the sample used and the methods of designing the instrument. Use of the instrument with at least two samples is required. The first application is to the so-called construction sample, which is the set of subjects used to develop the instrument. The intent usually is that the instrument reflect the original sample as accurately as possible under the assumption that the observed relations will be found also in future samples. This often results in a problem referred to as "sample overfitting"; it frequently is found that a "shrinkage" occurs in the magnitude of relations (for example, correlations or mean cost ratios) observed on samples other than the construction sample.

Cross validation is the only sure means of providing a check on the tendency to "overfit" samples. This requires the application of a developed instrument on another sample that is also representative of the target population. The resulting observed relations (for example, correlations or mean cost ratios) can indicate the likely predictive power that can be expected from the use of the instrument. Thus, validity coefficients, such as the mean cost ratio, found on validation samples, as compared to those on the construction sample, provide estimates of shrinkage. Most important, they give an estimate of the validity expected from the

instrument when it is used for prediction.

Sampling concerns

An underlying principle of the construction of risk assessment instruments is that both construction and validation samples must be representative of the target population. The latter is essential if generalizations are to be made to the target population. Nonrepresentative samples can be expected to reduce the validity of the instrument when used with new samples; the selected predictors may not necessarily be appropriate for the target population.

Besides the requirement of a representative sample, it is also necessary to have a sample that is sufficiently large. Small samples increase the probability that chance fluctuations will affect the results, thus increasing the error in the construction of the instrument.

Method of Combining Predictors

The construction of a risk assessment instrument requires the use of any of many available techniques to combine selected predictors so as to estimate expected outcomes. Much literature exists on a variety of methods used to accomplish this task; it outlines the arguments and evidence regarding the relative effectiveness of the various techniques.[14] Generally, the methods for combining pre-

dictors fall into three types: those that use predictors equally weighted; those that differentially weight the various predictor items; or, some configural approach.

E.W. Burgess, in 1928, developed the first parole prediction instrument.[15] Each predictor found related to the criterion was assigned one point regardless of its strength of association (or association with other items). The sum of the accumulated points determined the risk assessment score for each parolee or other subject. Each was weighted equally, which means that each had an equal contribution to the final score. Intercorrelations among the predictors are ignored when this method is used.

Currently, multiple regression or the linear discriminate function are commonly used to combine and weight predictors.[16] Theoretically, these are more sound methods for combining predictors, since account taken of any overlap that occurs among predictors (that is, intercorrelations) as well as of the relations of the criterion classification. The contributions of items to the coefficient of determination (R^2) provide a convenient method for selecting those predictors to be included in the risk screening instrument. The coefficient of determination measures the contribution of the predictors in combination to the "explanation" of variance in the criterion classifications.

Recent evidence has suggested that the power of the unweighted linear additive model (that is, the Burgess method)

is at least equal to that of multiple regression, and other methods commonly used.[17] Comparisons of methods have thus far indicated that no particular method of combining predictors has a strong advantage over any other methods. One such comparison of commonly used methods led to the conclusion that all techniques used resulted in virtually the same degree of predictive efficiency.[18] In this circumstance, it seems reasonable to use relatively simple, easily understood methods. The most effective method may be the one that is the easiest to implement in the field and, given the present "state of the art," this may be the Burgess model, which provides the simplest means for combining predictors. Alternative models, although more sophisticated, may be found to have less predictive utility because of shrinkage problems during the validation sample.

This brief review of methodological issues has outlined the major areas requiring attention when developing risk screening devices. It thus provides a foundation for examining and anlyzing the risk screening instruments employed by the ICFS project sites. The assessment of the utility of these risk screening devices requires attention, minimally and inter alia, to:

1) The representativeness of samples studied, in respect to the target population;
2) The operational definitions of the predictors and criterion;

3) The improvement in prediction using the instrument, compared with predictability from the base rate alone;
4) The validation of the instrument; and
5) The degree of validity found in operational use.

THE RISK SCREENING DEVICES

The agency personnel participating in the Improved Correctional Field Services Project used their own risk screening instruments in classifying their probationers. Each instrument was devised locally by either the research analyst of the participating agency or an outside consultant. Each can be reviewed in terms of the previously identified methodological issues, and this review should answer the question whether the risk screening instruments used are sufficiently valid for their intended purpose.

The ICFS program also provided the opportunity to seek to validate (or re-validate, in the case of Suffolk County) the risk screening instruments. The first cohort of probationers, a sample of 507 persons, included 102 from Kane County, 127 from Suffolk County, and 278 from Florida. This cohort consists of persons who had been on probation for at least six months at each agency. After a review of the methods used to develop each instrument, the validity of each will be assessed. The criterion that was employed for all the validation ICFS samples is re-arrest, reconviction, or revocation during the first six months of on-probation supervision. (A longer period of follow-up study, which would be desireable, is not yet possible.)

Kane County, Illinois

The risk screening instrument used in Kane County is

based on a linear, additive model, derived from a multiple regression analysis, but using arbitrary weights. The instrument includes eleven predictor items. Each has been assigned weights of one, two, or three, so the resulting scale is similar to that of Burgess. No assessment of reliability was reported, and the instrument was not validated before it was used in this project.

Construction Sample

The criterion employed for the construction sample was any recidivistic behavior of the probationers. Recidivism was defined as reconviction during the time frame of the study (15 months). Social demographic data items characterizing the probationers define the predictors.[19]

The construction sample consisted of 128 probationers who applied to the county's Community Correctional Services (CCS) program between January, 1976 and April, 1978. CCS operates in conjunction with the Sixteenth Judicial Circuit of Kane County as a broker for community services for probationers. The agency emphasis is on manpower services. Clients who enter the CCS program are either first time property offenders or eligible for Comprehensive Training and Employment Act (CETA) funding (the clients are generally unemployed). Clients eligible for CETA funding were used in the construction sample for the risk screening instrument. These clients provided a handy construction sample because CCS was the only division of the agency that routinely collected background information on probationers.

The data items collected by CCS were the only predictor item candidates considered for inclusion in the construction of the instrument. These were: sex, date of birth (age), marital status, racial group membership, type of Spanish-American (if applicable), student status, ability to speak English, migrant or seasonal worker, military status, handicapped, offender status, highest grade completed, economically disadvantaged, estimated annual income, receiving unemployment compensation, number of weeks on unemployment, labor force status, registered with ISES (the Illinois Employment Service), number in family, number of dependents, head of household, primary wage earner, most recent hourly wage, family receiving public assistance and type of assistance, CETA eligibility and current offender status (that is, diversion program, first-time, etc.). Items that referred to sex, age, or ethnic group membership were excluded. Correlations of the remaining items with the criterion were calculated; then, predictor items selected were dichotomized or categorized into three levels.

The selected outcome criterion, reconviction on probation, was determined from the date of application to CCS until the date the sample was collected. This means that

some probationers were at risk of failure longer than others. Indeed, the time on probation ranged from a minimum of six months (applied in April, 1978), to a maximum of 33 months (applied in January, 1976).

Of the 128 adult offender construction sample, 62 were in the failure category and 66 in the success category. The construction sample actually used, however, consisted of only 100 cases. Twenty-eight cases were excluded because they lacked a complete background file or outcome information. As already noted, the items used were those collected regularly by the CCS. Other items, such as previous criminal history, were thus not examined for possible use in the instrument.

The construction sample cannot be assumed to be representative of the target population. The CCS program is designed to serve a specific population, one which is not necessarily the same as those served by ICFS or the general population handled by the probation department. CCS clients have employment problems and/or are first time offenders. The intended ICFS target population was not restricted to persons with these characteristics. The exclusion of about a fifth of the original sample because of missing information also must be regarded as a source of potential bias.

After the criterion and predictors were selected, a stepwise regression program was carried out. The eighteen demographic variables were examined in terms of their con-

tribution to explaining the variance in the criterion. Any variable that contributed .005 percent or more to the coefficient of determination (R^2) was selected to be included in the instrument. The resulting model thus consists of the eleven predictors shown in Figure 1. The regression coefficients were not used in the assignment of weights; rather, as already indicated, arbitrary weights of 1, 2, or 3, were assigned to each predictor item category.

A risk score, which may range from 11 to 26, can be calculated for each probationer. Three risk categories were established on the basis of these scores: scores of 11-16 were categorized as high risk, 17-20 medium risk, and 21-26 low risk.

An examination of the results of the regression analysis, summarized in Table 1, provides additional information about the instrument. The original 18 variables had an $R^2 = .38$ in linear correlation with the criterion. This means that the variables in combination accounted for 38 percent of the variation in the dependent variable (reconviction or no reconviction on probation). Use of the .005 cut-off point rule eliminated seven variables, and it slightly reduced the "fit" with this sample. The decrease in predictors resulted in an $R^2 = .36$. The deletion of the seven variables thus appears completely warranted.

PRIMARY WAGE EARNER
1 = no
2 = yes

ESTIMATED ANNUAL INCOME
1 = $0-600
2 = $601-9,999
3 = $10,000 +

NUMBER OF WEEKS AT LONGEST JOB
1 = 0-11 weeks
2 = 12-104 weeks
3 = 105 + weeks

NUMBER OF WEEKS UNEMPLOYED
1 = 36 + weeks
2 = 3 - 35 weeks
3 = 0 - 2 weeks

ECONOMICALLY DISADVANTAGED
1 = no
2 = yes

REGISTERED WITH ISES
1 = no
2 = yes

MARITAL STATUS
1 = separated
2 = single, married, widowed
3 = divorced

NUMBER OF DEPENDENTS
1 = 0 - 2
2 = 3 +

HEAD OF HOUSEHOLD
1 = no
2 = yes

MOST RECENT HOURLY WAGE
1 = $0 - 2.64
2 = $2.65 +

DRIVER'S LICENSE
1 = no
2 = yes

TOTAL

Figure I
Kane County Risk Screening Instrument

As shown in Table 1, five variables accounted for around seventy percent of the explained variation (36 percent). They included: primary wage earner (eight percent), estimated annual income (six percent), number of weeks on longest job (four percent), economically disadvantaged (four percent), and registered with ISES (three percent).

Of the variables included in the risk assessment instrument, the zero order correlations with the outcome variable range from low to modest. Primary wage earner and number of weeks at longest job have the highest correlations with reconviction; .29 and .28 respectively. The items "possession of a driver's license" and "most recent hourly wage" clearly could be deleted without much loss, if any, in predictive efficiency. The intercorrelations among variables show that the predictors are not independent from one another. Table 2 displays these correlations for the 11 variables.

The efficiency of the instrument that might have been expected on the basis of the construction sample (given a similar population and about equal numbers of successes and failures) may be assessed by computing the mean cost rating and testing its statistical significance. The resulting values of the MCR for this instrument, using both the ungrouped scores and the operational score groupings, are reported in Table 3.[20]

TABLE 1
KANE COUNTY CONSTRUCTION SAMPLE
MULTIPLE REGRESSION WITH RECONVICTION ON PROBATION AS THE DEPENDENT VARIABLE

-PREDICTOR VARIABLE-	Zero Order Correlation (r)	Unstandardized Regression Weights (B)	Standardized Regression Weights (B)	Coefficient of Determination R^2	Variance Added R^2 Change
1. Primary Wage Earner	.286	.343	.336	.082	.082
2. Estimated Annual Income	.216	.293	.271	.142	.060
3. Weeks on Longest Job	.279	.152	.163	.186	.045
4. Weeks Unemployed	.204	.235	.300	.210	.023
5. Economically Disadvantaged	.060	.341	-.340	.247	.037
6. ISES Registered	.180	.176	.176	.277	.030
7. Marital Status	.135	.215	.224	.299	.022
8. Number of Dependents	.196	.352	.179	.317	.018
9. Head of Household	.141	-.341	-.340	.344	.027
10. Most Recent Hourly Wage	.208	.183	.123	.359	.014
11. Driver's License	.042	-.103	-.100	.364	.006

TABLE 2
KANE COUNTY CONSTRUCTION SAMPLE:
INTERCORRELATION OF PREDICTORS

		1	2	3	4	5	6	7	8	9	10	11
1.	Primary Wage Earner	1.0										
2.	Estimated Annual Income	- .10	1.0									
3.	Weeks on Longest Job	.16	.09	1.0								
4.	Weeks Unemployed	.06	.22	-.05	1.0							
5.	Economically Disadvantaged	.16	-.37	-.09	-.35	1.0						
6.	ISES Registered	.30	-.08	.05	.30	-.00	1.0					
7.	Marital Status	.02	-.04	.04	.06	-.14	.02	1.0				
8.	Number of Dependents	.14	.05	.13	.02	.09	.11	-.23	1.0			
9.	Head of Household	.82	-0.7	.09	.03	.11	.24	.02	.25	1.0		
10.	Most Recent Hourly Wage	.29	.13	.13	.02	-.17	.22	.01	.14	.82	1.0	
11.	Driver's License	.07	-.04	.12	.15	.08	.22	.01	-.12	-.05	-.04	1.0

TABLE 3
KANE COUNTY CONSTRUCTION SAMPLE:
PREDICTIVE EFFICIENCY OF THE INSTRUMENT

Risk Category	Raw Score	Outcome		Total
		Success	Failure	
HIGH	14	1	3	4
	15	0	0	0
	16	0	9	9
MEDIUM	17	3	12	15
	18	8	9	17
	19	14	8	22
	20	9	8	17
LOW	21	10	2	12
	22	7	0	7
	Total	52	51	103

Raw Scores (Uncollapsed Table)

MCR = .56

Z =4.88 p <.05

Collapsed Table

MCR = .43

Z =1.27 p >.05

The MCR for grouped scores (that is, when scores are grouped as done subsequently in the operational use of the instrument) is .43, which is not significant at the .05 level of confidence. The values of the MCR for the uncollapsed table (before risk scores have been grouped into categories), is .56, which is significant at the .05 level of confidence.

The MCR value for the ungrouped scores thus reflects a fairly substantial ability to discriminate subjects between the success and failure categories in the sample originally studied. Since, however, some information is lost when the probationers are classified into only these groups, the expected predictive efficiency is reduced markedly. And, it must be noted that the operational use of the instrument without a validation study can only be criticized as being unwise.

Validation Sample

The only validation sample available for the Kane County instrument is the ICFS first cohort used in the evaluation. The sample is drawn from a different population, but, after all, it was the ICFS eligible population with whom the instrument was expected to be used. The criterion definition is different; but it is the definition used for the ICFS evaluation study. It may

be noted that the base rate is markedly different.

The validation sample consisted of 101 ICFS project probationers who had been on probation for at least six months. (One case was deleted because the risk screening instrument did not have complete information.) Certain categories of probationers were excluded from the ICFS project. They included:

-Those sentenced to residential treatment (including treatment for drug abuse or for mental health problems) or to a work release program as a condition of probation;

-Those sentenced to a period of incarceration prior to being placed on probation;

-Those sentenced to less than six months of probation.

Thus, the sample studied cannot be assumed to be representative of the general probationer population handled by Kane County.

The assessment of validity of the instrument on the basis of its ability to discriminate favorable and unfavorable outcomes in the ICFS project sample is complicated by a possible effect of the classification/treatment process that provides the basis for the project and is the subject of the evaluation. That is, if the program is successful, the classification and differential treatment assignment will make a favorable difference; for example, if selection and more intensive supervision

of "high risk" probationers "works," then this risk is to a degree nullified. This is tantamount to saying that the expectation (from the risk instrument) is invalidated (as hoped). For further discussion and analysis of this issue, the reader must be referred to a companion report;[21] but this issue should be born in mind as a possible limitation of the analyses presented herein. (It should be noted that this circumstance obtains similarly in the case of ICFS validation samples for the other project sites as well).

A first step toward assessing the validity of the instrument was to repeat the step-wise regression with all eleven predicter items used. The coefficient of determination, as shown in Table 4, was .09. This means that these items, for this sample, account for only nine percent of the variation in the dependent variable. Thus it can be observed that a large shrinkage occurred (from $R^2 = .36$ to $R^2 = .09$). It is reasonable to account for this shrinkage by overfitting of the model to the construction sample, the use of samples from clearly different populations, and the use of a different criterion. The intercorrelations of items are shown in Table 5.

Most important, the validation sample demonstrates that the instrument is relatively inefficient in pre-

TABLE 4
KANE COUNTY ICFS VALIDATION:
MULTIPLE REGRESSION WITH ICFS OUTCOME AS THE DEPENDENT VARIABLE *

Offender Attributes	Zero Order Correlation	Standardized Regression Coefficient	Standard Error	Coefficient of determination (R^2)	Variance Added (R^2 Change)
Head of Household	.15	.450	.126	.021	.021
Driver's License	.14	.098	.101	.043	.022
Marital Status	.12	.359	.097	.052	.030
Registered with Employment Service	-.08	.262	.112	.067	.015
Number of Dependents	.12	.161	.139	.075	.008
Weeks Unemployed	.08	.167	.068	.081	.006
Wage Earner	.06	.088	.130	.083	.003
Economically Disadvantaged	.00	.059	.161	.087	.004
Weeks on Job	.10	.029	.074	.089	.002
Hourly Wage Earned		.152	.029	.090	.001
Annual Income	.05	.000	.090	.090	.000

*"Failure" is defined as rearrest, reconviction, or revocation during the first six months.

TABLE 5
KANE COUNTY ICFS VALIDATION SAMPLE:
INTERCORRELATION OF PREDICTORS

	1	2	3	4	5	6	7	8	9	10	11
1. Primary Wage Earner	1.00										
2. Estimated Annual Income	.15	1.00									
3. Number of Weeks at Longest Job	.10	.25	1.00								
4. Number of Weeks Unemployed	-.03	.27	.02	1.00							
5. Economically Disadvantaged	.26	-.37	-.11	-.45	1.00						
6. Registered with ISES	.10	-.08	.03	-.15	.18	1.00					
7. Marital Status	-.02	.00	.09	-.01	-.10	.27	1.00				
8. Number of Dependents	.15	.13	.31	-.08	-.02	-.01	-.10	1.00			
9. Head of Household	.68	.09	.13	-.06	.21	.19	.02	.15	1.00		
10. Most Recent Hourly Wage	-.02	.27	.07	.05	-.12	.07	-.07	-.03	-.02	1.00	
11. Driver's License	.07	.18	.14	.06	.03	-.17	.09	.20	-.03	.00	1.00

dicting, within the first six months of probation, rearrest, reconviction, or revocation. As Table 6 shows, the value of the MCR for the ungrouped scores is .26, which is not significant at the .05 level of confidence; for the present method of grouping, the value of MCR is .16, which is also not significant. It was noted previously that the construction sample and validation sample had different base rates for success. The construction sample had an intended base rate of .50, whereas the base rate of success for the validation sample was .74. This difference, along with the different criterion, sample, and short follow-up period of study, may help explain the instrument's poor predictive efficiency in the validation sample.

On this evidence, Kane County's risk prediction instrument does not meet the criterion of validity. The instrument does not provide an adequate basis for the risk classification used in the project.

Florida

The instrument used by the Salvation Army Department of Corrections in Florida was also constructed by using multiple regression techniques. The instrument consists of eight predictors: age, educational achievement, marital status, number of dependents, number of months employed, monthly income, prior probation or incarceration, and use of drugs or alcohol involved in the offense. (Figure II provides

TABLE 6

KANE COUNTY ICFS VALIDATION SAMPLE:
PREDICTIVE EFFICIENCY OF THE INSTRUMENT

Risk Category	Raw Score	Outcome		Total
		Failure	Success	
HIGH	14	1	2	3
	15	3	1	4
	16	4	7	11
MEDIUM	17	6	13	19
	18	5	11	16
	19	1	14	15
	20	3	14	17
LOW	21	2	6	8
	22	1	4	5
	23	1	1	2
	24	0	1	1
	25	0	1	1
	28			
	Total	29	75	102

Raw Scores (Uncollapsed Table)
MCR = .26
Z =1.93 p > .05

Collapsed Table
MCR = .16
Z =1.02 p > .05

AGE

1 = 20 or younger
2 = 21 - 27
3 = 28 +

EDUCATION

1 = 3 - 9 grades
2 = 10 - 11 grades
3 = 12 + grades

MARITAL STATUS

1 = single
2 = married, divorced, separated, widow(er)

NUMBER OF DEPENDENTS

1 = none
2 = one or more

MONTHS EMPLOYED

1 = none
2 = 1 - 12 months
3 = 13 + months

MONTHLY INCOME

1 = $0 - 199
2 = $200 - 599
3 = $600 +

PRIOR PROBATION OR INCARCERATION

1 = yes
2 = no

DRUGS/ALCOHOL INVOLVED IN COMMISSION

1 = yes
2 = no

TOTAL

Figure II

Florida Risk Screening Instrument

a copy of the instrument.) The criterion employed was rearrest or revocation with length of exposure to risk unknown (and hence, not taken into account in the analysis). This instrument was also not validated before it was implemented in the ICFS project. The validation sample consists of a sample of 278 probationers who participated in ICFS for a minimum of six months. Nine cases were deleted from the validation study because of missing information.

Construction Sample

The construction sample for the Florida instrument was 328 probationers who were terminated from probation during 1977 and early 1978 (through May). The sample was selected from a possible target population of 2,254 probationers from the three largest counties participating in the Salvation Army's misdemeanant program: Duvan County (Jacksonville), Dade County (Miami), and Pinellas County (Clearwater). This sample met the criteria for entry into the ICFS project in that none of the probationers were:

- residents in a treatment center for drugs, alcohol abuse, or other reasons;
- sentenced to work release or other residential program;
- sentenced to incarceration for more than 24 hours prior to being placed on probation;

- sentenced to probation for less than six months
- on simultaneous state or Federal probation.

The construction sample, thus, could be fairly representative of the target population; it should be noted, however, that the sample was chosen from urban areas of the state only; the rural areas were not included.

Of the probationers included in the sample, 102 were classified as failures and 226 as successes. The procedure for selecting the sample was as follows. All probationers who were either rearrested or reconvicted (n = 118) (the failures) and a 12 percent random sample of successful cases (255) were included. The sample size was decreased by 45 cases (16 failures and 29 successes) because they lacked sufficient data in their case files about months on present job or monthly income. (The cases excluded for lack of data on employment and income had a slightly lower (64 percent) success rate compared with the rest (69 percent). There may be some bias in this sample resulting from the exclusion of the unemployed, seasonal workers, or others with unstable employment.)

Several problems are inherent in this sampling procedure. First, the inclusion of a disproportionate number of failures increases the base rate over that expected in the population to which the instrument is to be applied.

The success/failure ratio for the original 2,254 probationers was about 18 successes for each failure, while the sample procedure resulted in about two successes for each failure. This difference in success/failure ratio could affect the selection of variables and the later application of the instrument. The ability to locate variables which discriminate between those cases likely to succeed or fail is increased when the base rate is made larger; but the problem of using the instrument to help identify relatively rare failures in the real world with a small base rate remains.

As in the case of Kane County, only a few data items were collected, reducing the likelihood of locating other variables that might distinguish among likely candidates in the success and failure categories.

For fourteen predictor item candidates, the correlations with revocation or rearrest while on probation were determined. The fourteen included: race, sex, age, marital status, number of dependents, education, employed full-time, receiving welfare, months on present job, monthly income, previous probation or incarceration, present offense, drugs or alcohol involved in the commission of the crime, and multiple offenses. From these, race and sex were eliminated as invidious. The multiple offenses item was also excluded without reported reasons.

The largest zero order correlation coefficients with

outcome were as follows: prior probation or incarceration (-.27), age (-.25), and months on present job and education (-.21). (The negative associations with the outcome result from the way the criterion was coded; they are in the expected directions).

The multiple regression analysis showed that the eight-variable instrument produced a multiple correlation coefficient of .40 with the outcome measure. The instrument thus accounts for 16 percent of the variance in the outcome for this sample. As shown in Table 7, the stepwise regression program used selected only five of the eight predictors, the fifth adding very little to the variance already explained. The other three variables did not contribute to the variance added in the coefficient of determination.

Table 7 suggests that the first four variables, in weighted linear combination, should provide as much predictive efficiency as the eight item scale. The intercorrelations of these items are shown in Table 8. For the most part, the variables are relatively independent, with the intercorrelations not exceeding .20. Age has several modest correlations with marital status (.52), number of dependents (.38), months employed on present job (.25), and monthly income (.17). Substantial correlations appear only between the number of dependents and marital status (.73).

TABLE 7
FLORIDA CONSTRUCTION SAMPLE:
MULTIPLE REGRESSION WITH "FAILURE" AS THE DEPENDENT VARIABLE *

	Zero Order Correlations	Unstandardized Regression Weights	Standardized Regression Weights	R^2	R^2 Change
Prior Incarceration or Probation	-.27	-.23	-.22	.07	.07
Age	-.25	$-.96^{01}$	-.15	.12	.05
Months on Job	-.21	$-.89^{01}$	-.16	.14	.02
Education	-.21	$-.84^{01}$	-.13	.16	.02
Drugs/Alcohol Use	-.10	$-.71^{01}$	-.07	.16	.005

* "Failure" was defined by rearrest or reconviction.

TABLE 8
FLORIDA CONSTRUCTION SAMPLE:
INTERCORRELATION OF PREDICTORS

	1	2	3	4	5	6	7	8
1. Age	1.00							
2. Education	.19	1.00						
3. Months on Job	.25	.10	1.00					
4. Prior Incarceration or Probation	.11	.14	.05	1.00				
5. Drug/Alcohol Use	.09	.05	-.11	.13	1.00			
6. Marital Status	.52	-.06	.12	.09	.15	1.00		
7. Monthly Income	.17	.05	.19	-.01	-.02	.11	1.00	
8. Number of Dependents	.38	-.12	.09	.08	.20	.73	.02	1.00

The regression weights were not used; rather, items were scored as shown in Figure II, providing, as in Kane County, a modified Burgess-type instrument. The relative efficiency of the instrument in the sample studied may be assessed from the data shown in Table 9. The mean cost rating for the grouped scores is .34, which is significant at the .001 level. A comparison with the ungrouped scores suggests that the currently used groupings do result in some loss of information (MCR = .42). That the MCR is statistically significant gives some credence to the instrument; but this must be viewed cautiously in view of the lack of validation, a problem that is important particularly in the light of the problems with potential sample bias and the base rate. Confidence in the use of this instrument could come only from evidence from validation studies on samples representative of the population to be included in the ICFS project.

Validation Sample

The ICFS project evaluation provided an opportunity, as in the other two sites, for a limited assessment of the validity of the instrument used in Florida. It should be noted again that a somewhat different criterion was used.

The validation sample consisted of 269 probationers who had been on probation for a minimum of six months.

TABLE 9
FLORIDA CONSTRUCTION SAMPLE:
PREDICTIVE EFFICIENCY OF THE INSTRUMENT

Category	Score	Outcome: Success	Outcome: Failure	Total
HIGH	8	0	1	1
	9	0	2	2
	10	3	8	11
	11	3	8	11
MEDIUM	12	21	13	34
	13	21	17	38
	14	18	13	31
	15	29	9	38
	16	35	14	49
LOW	17	26	8	34
	18	50	8	58
	19	11	1	12
	20	9	0	9
	Total	226	102	328

Raw Scores (Uncollapsed Table)
MCR = .42
Z =6.08 $p < .05$

Collapsed Table
MCR = .34
Z =5.78 $p < .05$

The sample was drawn from the 13 district offices of the Salvation Army Department of Corrections participating in the ICFS project. All probationers who met the project inclusion criteria were included.

A stepwise multiple regression analysis was completed and the results are summarized in Table 10. The coefficient of determination, .07, suggests that these items in combination have little predictive utility for this new sample. The shrinkage from .16 to .07 could be due largely to the method used to select the construction sample. The different criterion used could also contribute to the shrinkage in the amount of variance explained.

As Table 10 illustrates, the zero order correlation coefficients show very low associations between the predictors and the criterion. Both education and drugs-alcohol involved in the offense have the largest correlation with the criterion; but it is only .14.

The intercorrelations of predictors in the validation sample are shown in Table 11. The same overlapping relations appear in the validation sample as in the construction sample: age and marital staus, .61; number of dependents and marital status, .56; age and number of dependents, .39; and months employed and monthly income, .50.

The value of the MCR's as reported in Table 12 for the uncollapsed and collapsed tables are .20 and .15, respectively. The MCR for the uncollapsed table is significant at the .05 level; the MCR for the collapsed version

TABLE 10
FLORIDA ICFS VALIDATION:
MULTIPLE REGRESSION WITH ICFS OUTCOME AS THE DEPENDENT VARIABLE *

Offender Attribute	Zero Order Correlation With Outcome	Standardized Regression Coefficient	Standard Error	Coefficient of Determination (R^2)	Variance Added (R^2 -Change)
Education	.140	.588	.030	.022	.022
Age	.121	.698	.035	.041	.019
Prior Probation	.124	.289	.048	.054	.013
Months Employed	-.030	.217	.031	.062	.008
Drugs/Alcohol on Day of Offense	.136	.176	.047	.066	.004
Dependents	.006	.098	.051	.070	.000
Monthly Income	.03	.011	.033	.070	.000
Marital Status	.04	.002	.059	.070	.000

* "Failure" is defined as rearrest, reconviction, or revocation during the first six months.

TABLE 11
FLORIDA ICFS VALIDATION SAMPLE:
INTERCORRELATION OF PREDICTORS

	1	2	3	4	5	6	7	8
1. Age	1.00							
2. Education	.03	1.00						
3. Marital Status	.61	-.04	1.00					
4. Number of Dependents	.39	-.05	.56	1.00				
5. Months Employed	.23	.17	.17	.03	1.00			
6. Monthly Income	.27	.10	.29	.16	.50	1.00		
7. Prior Probation or Incarceration	-.13	.11	-.10	-.07	.05	.05	1.00	
8. Drugs/Alcohol Involved in Commission	.08	.06	.09	.88	-.07	.07	.26	1.00

TABLE 12
FLORIDA ICFS VALIDATION SAMPLE:
PREDICTIVE EFFICIENCY OF THE INSTRUMENT

RISK CATEGORY	RAW SCORES	OUTCOME		TOTAL
		Failure	Success	
HIGH	8	1	1	2
	9	3	1	4
	10	1	4	5
	11	2	13	15
MEDIUM	12	7	30	37
	13	3	29	32
	14	6	30	36
	15	6	31	37
	16	2	32	34
LOW	17	2	24	26
	18	2	15	17
	19	2	17	19
	20	1	4	5
	TOTAL	38	231	269

Raw Scores (Uncollapsed Table)
MCR = .20
Z = 1.96 $p < .05$

Collapsed Table
MCR = .15
Z = 1.55 $p > .05$

of the instrument is not significant. The MCR's are considerably lower for the validation sample than for the construction sample. It is justifiable to conclude that for the ICFS project clientele, this instrument has a very low validity and can not be expected to be very useful when used as in the ICFS project -- that is, with the information loss inherent in classification of all probationers into three groups.

In summary, the instrument used in Florida has little predictive utility and did not provide an adequate risk classification. The instrument was devised without considering the base rate problem. The method used to select the sample limited the degree to which the sample could be considered representative of the target population (ICFS project probationers). And, the instrument was used with no assessment of its validity. The scale has a low degree of validity, but that is reduced even further when used to simply group probationers into three groups.

Suffolk County

Unlike the other two risk screening devices, the Suffolk County instrument focuses on two aspects of the probationer's characteristics: "the offender's involvement on a cluster of major variables" and "the severity

level on particular variables".[22] The Suffolk County instrument, entitled Differential Classification Form for the Supervision of Probationers (Figure III), reflects the agency's concern with areas of major dysfunctions that could result in further criminal activity. This instrument is thus intended to be used as a combination risk and needs assessment screening device.

The instrument is designed to assess the probationer's status in certain key areas, and the degree of "severity" of this status. These key components define the two scales of the instrument. The first scale refers to an individual's status in four major areas: the nature of the current offense, prior evidence of psychological instability, any prior record within the past seven years, and evidence of social instability. Each "status" condition is worth one point; thus, a maximum of four points can be assigned on the status scale.

On the second scale, the severity of the status problem is assessed. Under each status condition there are several items that determine the nature and severity of the status or dysfunction. For instance, within the psychological instability status the items are: hospitalization or commitment to a rehabilitative

SUFFOLK COUNTY DEPARTMENT OF PROBATION

DIFFERENTIAL CLASSIFICATION FORM FOR THE SUPERVISION OF PROBATIONERS

NAME: CASE NO: DATE:

PROBATION OFFICER: S.P.O: OFFICE:

A. Current Offense - Status: Yes ☐ No ☐ SCORE
1. Felony conviction(s) (2 pts) ☐
2. Assaultive conviction(s) (2 pts) ☐
3. Driving While Intoxicated (1 pt). ☐

B. Psychological Instability: Yes ☐ No ☐
1. Hospitalization or committment to a rehabilitative program (2 pts) ☐
2. Diagnosed psychotic, severely emotionally disturbed, severely retarded (2 pts) ☐
3. Alcohol or drug dependent (2 pts) ☐

C. Prior Record (last 7 years) : Yes ☐ No ☐
1. Felony conviction(s) (2 pts) ☐
2. Misdemeanant conviction(s) (2 pts) ☐
3. Youthful Offender (YO) convictions (2 pts) ☐
4. Juvenile Delinquency (JD) Adjudication (1 pt) ☐

D. Social Instability: Yes ☐ No ☐
1. Educational vocational, employment deficits (1 pt) ☐
2. Weak, non-existent positive family or community attachments (1 pt) ☐
3. Recidivism or Recidivistic tendencies (2 pts) ☐

E. Age: This variable is only used for marginal cases. Between 16 - 24 years old (1 pt) ☐

Variables	Status	Level of Severity
A. Current Offense		
B. Psychological Instability		
C. Prior Record		
D. Social Instability		
Subtotal		
E. Age	XXXXX	
Total		

LENGTH OF SENTENCE: TIME SERVED: LENGTH OF A.S:

CLASSIFICATION: Intensive ☐ Active ☐ Special ☐

Figure III
Suffolk County Risk Screening Instrument

program, diagnosed psychotic, severely emotionally disturbed, and alcohol or drug dependent. Each "severity" variable has a value ranging from 1 to 2 points. If a person is between 16-24 years old, another point can be added on the severity scale.

Both scales are used to determine the person's risk level. The scales are added separately and then joined according to the following formula: risk = status + (.01) severity. All cases with 0 points are assigned low risk, 1.01 to 3.06 medium, and 3.07 or above high risk. It is predicted that seven percent of the low risk, 29 percent of the medium risk, and 74 percent of the high risk clients will fail.[23]

Construction Sample

In 1974 the Suffolk County Department of Probation, investigating the feasibility of a classification system, undertook an analysis of approximately 2.2 percent of their 33,250 probation cases (N = 720 adult probationers). Case records were examined and the following data items were extracted from the files: prior criminal history, degree of community attachments, marital status, and alcohol use and abuse. (Some additional items that could be found in the presentence investigation were also included.) The development of a classification device proceeded with the identification

of areas of perceived major behavioral dysfunctions for the probation population.

A "backdoor" configural approach was used to analyze the data and frame the classification instrument.[24] The criterion used was reconviction during probation or after probation termination. The time frame of the follow-up is unknown. The length of time at risk varied for individuals in the sample (and was not taken into account). About 25-30 percent of the sample was classified as failures. The resulting instrument was derived by examining how factors appeared to cluster. From this anlysis, two levels of supervision were instituted; Intensive and Normal.

The first draft of the instrument was tested in a pilot study conducted during the following year. This study was based on 627 persons. The pilot study provided an opportunity to validate the instrument, to clarify the design, to obtain feedback from the probation officers, and to determine the degree of agreement between probation officers' recommendations for supervision level and those made on the basis of the instrument.

This pilot study was conducted in two phases. The initial classification phase required the probation officers who were supervising each case to classify these cases according to the objective criteria of the differential classification system. The second phase of this

> study compared the subjective classification system of each probation officer against the objective criteria .... After all cases had been subjectively and objectively classified, the probation officers were asked to compare the results and explain the contradictions.[25]

With a 95 percent agreement between the probation officers' subjective classifications and the formal, objective recommendation for classifying individuals by the instrument, only a few minor refinements were necessary. The revised instrument was then implemented. Periodically, the operational definitions of the items were changed to accommodate new characteristics of the probationers. For instance, the definition of time frames for hospitalization and prior record were changed to reflect closer proximity to the date placed on probation for the current offense.

The development of this instrument differs markedly from that in the other two ICFS sites. The first phase involved the construction of the instrument on the basis of study of a random sample of the available adult population. This ensured a sample that was representative of the population at large; but it did not necessarily assure a sample representative of the ICFS target population.

Initially, there appears to have been some confounding of the risk and needs assessment objectives.

In the early stage of development, reconviction while on probation or post-probation was used as the criterion; during the pilot study the criterion became the degree of agreement between classifications by the officers and by the instrument as to the appropriate level of supervision. Further validation study relative to the risk assessment aim was accomplished, however, as next described.

Validation Sample

The Suffolk County Probation Department completed a validation study of their instrument in order to determine whether to use it in the ICFS project. (Use of an existing New York State Intensive Supervision Project risk instrument was an available alternative.) The Suffolk County instrument was adopted because the association (that is, the obtained value of Chi square) was slightly higher with the Suffolk instrument than with the ISP device.[26] The criterion of failure used in this comparison validation was rearrest while on probation or during three years following termination of probation. The validation sample consisted of 355 cases, including, it should be noted, probationers sentenced to a residential program, having a split sentence, and who were convicted on a felony sex offense. (These probationers had to be excluded from the ICFS project according to the ICFS

guidelines.)

Using these data, a series of regression analyses was performed by the Suffolk County staff. The purpose was to determine the predictive efficiency of the instrument, to examine the intercorrelations of predictors, and to examine the instrument from a different framework than that of its original design. The stepwise regression analysis of the Suffolk validation study provides additional information for the refinement of the instrument.

The items in the Suffolk County instrument, based on a sample of 351 cases (four cases were deleted because of missing values), explain about 38 percent of the variance in probation outcomes (Table 13). As typically found in analyses of this sort, most of the contribution to the outcome variance is accounted for by the first few items selected. In this case, most of the contribution is made by the first five or six items (at most). A similar result is shown in Table 14 in which the analysis was limited to the "severity" scores.

The correlation coefficients between the items and the outcome criterion are shown also in Tables 13 and 14. The largest zero order correlations are: recidivism or officers' perception of recidivistic tendencies (.48), age (.41), prior criminal record

TABLE 13
SUFFOLK COUNTY VALIDATION SAMPLE:
MULTIPLE REGRESSION WITH RECONVICTION AS THE DEPENDENT VARIABLE AND ALL VARIABLES AS PREDICTORS

	Zero Order Correlation	Unstandardized Regression Wgt.	Standardized Regression Wgt.	R^2	R^2 Change
Recidivistic Tendencies	.48	.18	.36	.23	.23
Age	.41	.19	.18	.30	.07
Prior Record Status	.35	.28	.31	.33	.03
Weak Ties	.30	.15	.14	.34	.02
Alcohol/Drug Dependence	.20	.13	.23	.36	.01
Misdemeanor Convictions	.16	$-.92^{01}$	-.19	.37	.01
Educational Deficits	.27	.11	.13	.37	.004
Psychological Instability Status	.17	$-.86^{01}$	-.08	.38	.003
Social Instability Status	.32	$-.99^{01}$	-.10	.38	.003
Youthful Offender Status	.28	$-.42^{01}$	-.06	.38	.002
Driving While Intoxicated	-.04	$-.96^{01}$	-.07	.38	.002
Hospitalization	.15	$-.22^{01}$	-.03	.38	.000
Current Offender Status	.04	$-.30^{01}$	.03	.38	.000
Psychotic	.03	-.24	-.02	.38	.000
Assaultive Conviction	-.02	$-.99^{02}$	-.01	.38	.000
Prior Felony Conviction	.18	$.12^{01}$	.01	.38	.000
Juvenile Delinquent	.23	$.20^{01}$	.01	.38	.000

TABLE 14
SUFFOLK COUNTY VALIDATION SAMPLE
MULTIPLE REGRESSION WITH "SEVERITY" VARIABLES AS THE INDEPENDENT VARIABLES AND RECONVICTION AS THE DEPENDENT VARIABLE

	Zero Order Correlation	Unstandardized Regression Weight	Standardized Regression Weight	R^2	R^2 Change
Recidivistic Tendencies	.48	.15	.31	.12	.23
Age	.41	.22	.20	.30	.07
Alcohol/Drug Dependence	.20	$.97^{01}$	.17	.32	.02
Weak Ties	.30	.13	.12	.34	.02
Juvenile Adjudication	.23	.14	.09	.35	.007
Educational Deficits	.27	$.61^{01}$	.06	.35	.004
Hospitalization	.15	$-.41^{01}$	-.06	.35	.003
Youthful Offender Convictions	.28	$-.29^{01}$	.04	.35	.001
Prior Felony Convictions	.18	$.27^{01}$	.3	.36	.001
Driving While Intoxicated	-.04	$-.57^{01}$	-.04	.36	.001
Psychotic	.03	$-.47^{01}$	-.04	.36	.001
Misdemeanor Convictions	.16	$.19^{01}$	.04	.36	.001
Current Felony Convictions	.10	$.83^{02}$	.01	.36	.001

status (.35), social instability status (.32), weak community or family ties (30), youthful offender conviction (.28), and education or vocational deficits (.27). The remaining variables have a low or negligible correlation with the outcome variable.

The examination of the intercorrelations of predictor items tends to support the a priori "clusters" of the scales' originators. As seen in Table 15, current offense status is correlated substantially with felony conviction, assaultive conviction, and driving while intoxicated. Similar moderate to substantial correlations are found for "severity" scores and the other three "status" items. The status variables do not appear to be correlated substantially with any other predictors, which also supports the theory of the scale developers. Few of the "severity" predictors appear to be intercorrelated except for age and youthful offender conviction (.42), age and recidivism (.33), educational deficits and weak community ties (.34), alcohol/drug dependent and hospitalization (.46), and psychotic and hospitalization (.39).

The Suffolk County instrument exhibits in this sample a substantial predictive capacity. As shown in Table 16, the values of the MCR for the uncollapsed and collapsed version of the screening device are .56

TABLE 15
SUFFOLK COUNTY VALIDATION SAMPLE:
INTERCORRELATION OF PREDICTORS

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1. CURRENT OFFENSE STATUS	1.00																	
2. Felony Conviction	.51	1.00																
3. Assaultive Conviction	.45	.13	1.00															
4. DWI	.55	-.10	-.00	1.00														
5. PSYCHOLOGICAL INSTABILITY	.23	.02	.04	.32	1.00													
6. Hospitalization	.07	.06	.16	.06	.63	1.0												
7. Psychotic	.03	.04	-.08	0.04	.37	.39	1.0											
8. Alcohol/Drug Dependent	.23	.03	-.01	.36	.87	.46	.16	1.0										
9. PRIOR RECORD	.15	.04	.03	.15	.26	.16	.13	.24	1.0									
10. Felony Conviction	.07	.17	-.04	-.04	.24	.22	.19	.22	.27	1.0								
11. Misdemeanant Conviction	.16	-.02	.10	.26	.34	.17	.16	.33	.76	.28	1.0							
12. Y.O. Conviction	.04	.15	.01	-.14	.03	.10	.09	.05	.40	.08	.05	1.0						
13. Juvenile Adjudication	.02	.09	-.11	-.07	.02	.11	.04	.03	.30	.05	.01	.17	1.0					
14. SOCIAL INSTABILITY	-.06	.02	-.07	-.05	.13	.18	.07	.08	.13	.12	.10	.16	.10	1.0				
15. Educational Deficits	.01	.09	-.07	-.04	.16	.15	.04	.13	.17	.12	.11	.14	.07	.74	1.0			
16. Weak Ties	-.06	.09	-.01	-.13	.19	.29	.15	.10	.18	.15	.08	.22	.13	.45	.34	1.0		
17. Recidivism	-.00	.05	.08	-.00	.07	.15	.07	.09	.26	.18	.16	.30	.19	.55	.25	.25	1.0	
18. AGE (16 - 24 YEARS OLD)	.03	.17	.16	-.13	.07	.13	-01	.05	.33	.14	.10	.42	.22	.26	.23	.25	.33	1.0

TABLE 16
SUFFOLK COUNTY VALIDATION SAMPLE:
PREDICTIVE EFFICIENCY OF THE INSTRUMENT

Risk Category	Raw Score	Outcome: Success	Outcome: Failure	Total
Low	0	54	3	57
Medium	1.01	25	4	29
	1.02	38	6	44
	1.03	16	5	21
	1.04	2	3	5
	1.05	1	2	3
	2.02	2	0	2
	2.03	29	3	32
	2.04	20	5	25
	2.05	10	5	15
	2.06	5	10	15
	2.07	3	7	10
	2.08	4	3	7
	2.09	1	3	4
	3.04	4	0	4
	3.05	10	3	13
	3.06	6	4	10
High	3.07	3	8	11
	3.08	4	4	8
	3.09	2	2	4
	3.10	1	3	4
	3.11	1	2	3
	3.12	1	3	4
	3.13	0	1	1
	3.15	0	1	1
	4.07	0	2	2
	4.08	2	1	3
	4.09	1	1	2
	4.10	0	2	2
	4.12	0	2	2
	4.13	0	1	1
	4.15	0	1	1
	4.16	0	2	2
	4.17	1	1	2
	4.18	1	1	2
	Total	247	104	351

Raw Scores (Uncollapsed Tables) MCR = .56 Z=8.35 $p < .05$
Collapsed Tables MCR = .41 Z=6.13 $p < .05$

and .41, respectively, both statistically significant, and higher than values usually obtained with similar risk instruments.

Revalidation

The ICFS sample of 127 probationers was used to revalidate the Suffolk County instrument. This validation sample is useful in determining how well the instrument functioned in the ICFS project evaluation, bearing in mind the different natures of the samples studied and of the outcome criterion, and noting also that it is the ICFS eligible population with whom the instrument is used operationally for differential case assignment.

The validation sample consisted of 127 probationers; three cases were excluded because of missing values on some items of the risk screening instrument. The project, as already noted, excluded cases with sentences to a placement in a residential facility for treatment, with several convictions for driving while intoxicated, and with a period of incarceration prior to the probation sentence. The base rate of failures was .18, which is lower than that used in the Suffolk validation sample. But of course the follow-up period was limited to only six months on probation, while the earlier validation sample had a maximum of three years (both on or off probation).

The stepwise regression analysis shows that the

coefficient of determination is .27. This means that 27 percent of the variance in the outcome measure is explained by the items included. As shown in Table 17, the first seven variables account for 90 percent of the explained variance. The remaining eleven variables did not contribute much additionally. Table 17 shows the summary of the stepwise regression analysis excluding the status variables. The coefficient of determination for the severity variable model is .26; that is about the same as the model including both the status and severity variables. This suggests that the status variables do not add much.

The correlations between the predictors and the outcome variable tend to be small. Tables 17 and 18 show these correlations to range in value from ± .01 to .22. These are low associations with the outcome measure. Noticeable decreases occurred in the association between the predictors and outcome in this revalidation sample compared to the original validation sample.

The overlapping relations of severity and status variables (the originators' a priori "clusters") is evident in the ICFS sample as well as in the original validation sample. Table 19 shows the correlation coefficients of variables in the risk screening device. The highest correlations occur within status "dimensions". Remaining correlations tend not to be so high as

TABLE 17

SUFFOLK COUNTY ICFS VALIDATION:
ICFS OUTCOME AS THE DEPENDENT VARIABLE AND MULTIPLE REGRESSION WITH ALL VARIABLES AS PREDICTORS

OFFENDER ATTRIBUTES	Zero Order Correlation w/ Outcome	Standardized Regression Coefficient	Standard Error	Coefficient of Determination (R^2)	Variance Added R^2 Change
Alcohol/ Drug Dependence	-.219	.493	.208	.058	.058
Driving While Intoxicated	.152	1.57	.198	.136	.078
Recidivistic Tendencies	-.134	.486	.100	.173	.095
Prior Record Status	-.127	.011	.352	.210	.027
Felony Conviction	.192	.459	.057	.231	.021
Age	.042	.263	.427	.246	.014
Psychotic	.030	.213	.184	.250	.004
Psychological Instability Status	-.163	.128	.377	.254	.004
Current Offense Status	.103	.108	.129	.258	.004
Misdemeanor Convictions	-.055	.125	.188	.262	.004
Hospitalization	-.073	.045	.162	.265	.003
Assaultive Conviction	-.013	.046	.056	.267	.002
Juvenile Delinquent	-.105	.075	.438	.268	.001
Prior Felony Conviction	.040	.061	.526	.269	.001
Youthful Offender	-.088	.057	.193	.272	.003
Educational Deficits	-.192	.000	.123	.272	.003
Weak Ties	-.133	.000	.081	.272	.003
Social Instability Status	-.213	.000	.146	.272	.003

TABLE 18

SUFFOLK COUNTY ICFS VALIDATION:
MULTIPLE REGRESSION WITH ICFS OUTCOME AS THE DEPENDENT VARIABLE * AND ALL "SEVERITY" VARIABLES AS PREDICTORS

OFFENDER ATTRIBUTES	Zero Order Correlation w/ Outcome	Standardized Regression Coefficient	Standard Error	Coefficient of Determination (R^2)	Variance Added R^2 Change
Alcohol/ Drug Dependence	-.219	1.10	.072	.058	.058
Driving While Intoxicated	.152	1.75	.146	.136	.078
Recidivistic Tendencies	-.134	.826	.077	.173	.095
Felony Conviction	.192	.342	.044	.207	.024
Misdemeanor Conviction	-.055	.421	.063	.221	.014
Age	.042	.232	.414	.239	.018
Juvenile Delinquent	-.105	.139	.256	.246	.007
Youthful Offender	-.088	.128	.081	.250	.004
Prior Felony Conviction	.040	.072	.385	.254	.004
Psychotic	.030	.092	.074	.258	.004
Hospitalization	-.073	.048	.159	.260	.002
Educational Deficits	-.192	.006	.070	.261	.001
Weak Ties	-.133	.002	.077	.261	.000
Assaultive Convictions	-.013	.001	.040	.261	.000

* "Failure" is defined as rearrest, reconviction, or revocation during the first six months.

TABLE 19
SUFFOLK COUNTY ICFS VALIDATION SAMPLE:
INTERCORRELATION OF PREDICTORS

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1. CURRENT OFFENSE STATUS	1.0																	
2. Felony Conviction	.52	1.00																
3. Assaultive Conviction	.67	.36	1.00															
4. DWI	.35	-.10	-.15	1.00														
5. PSYCHOLOGICAL INSTABILITY	.16	-.03	.03	.35	1.00													
6. Hospitalization	.17	-.02	.05	.28	.67	1.00												
7. Psychotic	.15	.18	.27	.03	.48	.26	1.00											
8. Alcohol/Drug Dependent	.09	-.13	-.12	.44	.79	.60	-.10	1.00										
9. PRIOR RECORD	.14	-.26	-.01	.45	.35	.34	.03	.36	1.00									
10. Prior Felony Conviction	-.08	-.05	-.06	-.03	-.04	-.03	-.03	-.03	-.05	1.00								
11. Misdemeanant Conviction	.15	-.20	-.10	.48	.43	.37	.09	.41	.79	-.04	1.00							
12. Y.O. Conviction	-.05	-.11	0.2	-.07	.00	.06	-.06	.04	.36	.40	-.09	1.00						
13. Juvenile Adjudication	.14	-.06	.20	-.04	-.06	-.04	-.03	-.05	.26	-.01	-.05	-.03	1.00					
14. SOCIAL INSTABILITY	.04	-.12	.14	-.05	.11	.05	-.02	.18	.07	.06	.09	.08	-.04	1.00				
15. Educational Deficits	-.10	-.18	.00	-.13	.05	.01	-.15	.16	.10	.01	.09	.13	-.01	.78	1.00			
16. Weak Ties	.14	-.10	.07	.02	.15	.22	-.13	.27	.07	.16	.12	.06	-.06	.36	.32	1.00		
17. Recidivism	.08	.09	.10	.08	.00	-.03	.09	-.02	-.03	-.05	-.05	-.03	.08	.40	.02	.09	1.00	
18. AGE (16-24 YEARS OLD)	.10	-.05	.14	-.03	.19	.29	.34	-.03	.18		.23	-.02	-.01	.07	-.10	-.05	.16	1.00

within these status groups.

The predictive efficiency shown in the original validation sample (with respect to Suffolk probationers generally) was not repeated in the revalidation (concerning ICFS project probationers). The different criterion (and, particularly, the markedly different follow-up periods) should be remembered. The MCR's are .20 for the ungrouped risk scores and .10 for the risk classifications, as shown in Table 20. Neither MCR is statistically significant.

ICFS Risk Screening Instruments

Each of the three risk screening instruments employed by the ICFS sites had some methodological shortcomings. Two of the instruments, in Kane County and Florida, suffered from a lack of any validation: a problem enhanced by a lack of attention to the base rate problem. None of the instruments was validated on a sample that could be assumed to be representative of the population to which the instruments would be applied. These problems are sufficient to undermine the integrity of implementation of the general design of the ICFS project -- differential supervision by risk (as determined by an appropriate instrument). The risk classifications did not have demonstrable validity for the target population at any site at the start of the project. An additional objective of the

TABLE 20
SUFFOLK COUNTY ICFS VALIDATION SAMPLE:
PREDICTIVE EFFICIENCY OF THE INSTRUMENT

Risk Category	Raw Score	Outcome 0	Outcome 1	Total
LOW	0	3	19	22
MEDIUM	1.01	1	14	15
	1.02	2	11	13
	1.03	2	7	9
	1.04	1	8	9
	2.03	4	16	20
	2.04	2	7	9
	2.05	0	4	4
	2.06	1	3	4
	2.07	1	2	3
	2.08	0	1	1
	3.05	1	2	3
	3.06	2	0	2
HIGH	3.07	1	1	2
	3.08	1	2	3
	3.1	0	1	1
	4.04	0	1	1
	4.07	0	1	1
	4.10	1	1	2
	Total	23	101	124

Raw Score (Uncollapsed Table)
MCR = .20
Z = .01 $p > .05$

Collapsed Table
MCR = .10
Z = .48 $p > .05$

analyses reported here, besides contributing to the evaluation of the overall project, has been to provide validation information by using the project samples.

Further Explorations of Risk

In order to possibly contribute to improving risk screening at the sites, but also to achieve other objectives of the ICFS evaluation as well, the data from the first six months of experience in the project were further analyzed to create an additional measure of risk. (As described in a companion report, the new risk measure described in this section was used as a statistical control for possible biasing factors in the classifications into supervision categories).

The existing instruments are restricted in their use of predictors. For the most part, these predictors are social demographic characteristics of the probationers. In item selection, prior criminal history variables that have demonstrated in other studies their usefulness in determining risk level, were neglected. The ICFS project evaluation staff coded 87 potential predictors. The following items are examples: number of prior adult arrests, age at first arrest, length of residence, number of changes in address during last year, substance abuse, and type of prior criminal history. The construction of a new

risk screening instrument will focus on using as prediction candidates a variety of offender attributes that were not necessarily included in the instruments already discussed.

In order to identify items, a multiple regression analysis was completed, with the 506 ICFS subjects (from the three sites combined) serving as the construction sample. For each probationer, background and follow-up data were coded from the existing case files.

The criterion used (as in the "revalidation" studies already described) was recidivism during the first six months under probation supervision. Recidivism includes rearrest, reconviction, or revocation during this period. Eighteen percent of the ICFS probationrs fell into this failure category.

Possible predictors were selected in two steps. First, all variables were examined to see whether they were associated with the outcome measure. Chi-square or Pearson's correlation coefficient was used, depending on the level of measurement of the attribute or variable, to determine the statistical significance of the association. The .05 level of confidence was used as the decision rule to accept a variable as being significant. Selected interval level

variables then were grouped (for simplicity, though at a possible cost of loss of some information); and, Chi square tests were applied with the resulting categories to test for significant associations with outcome. Next, the items that passed this test were included in a stepwise multiple regression analysis. Thirteen variables which in combination produced the highest coefficient of determination (that is, the squared multiple correlation coefficient) were retained, although as usual only the first few items selected account for most of the outcome variance "explained." The multiple correlation coefficient, using all items, was only .35, so the coefficient of determination was a quite modest .12. A summary of the analysis is shown in Table 21. None of the items had a strong correlation with the outcome, but all zero order correlations, while low, were statistically significant at the five percent level of confidence.

The correlation coefficients in Table 22 show that the highest intercorrelation is between prior arrest for a property offense and conviction for the same (-.78). Other variables that were substantially correlated are: age at first arrest and age at probation sentence (-.61), age at first conviction and age at the current probation sentence (.51), and substance abuse and drug abuse (.62).

TABLE 21

COMBINED SAMPLES:
MULTIPLE REGRESSION WITH ICFS OUTCOME AS THE DEPENDENT VARIABLE *

OFFENDER ATTRIBUTE	ZERO ORDER CORRELATION WITH OUTCOME	STANDARDIZED REGRESSION COEFFICIENT	STANDARD ERROR	COEFFICIENT OF DETERMINATION	VARIANCE ADDED (R^2 CHANGE)
Age at First Conviction	.193	.19	.82	.037	.037
Prior Adult Arrests	.178	.46	.53	.061	.024
Educational Status	.167	.89	.38	.076	.015
Age at Time of Probation	.146	.95	.37	.086	.010
Drug Abuse Problem	.166	.19	.82	.094	.008
Nature of Current Charge	.094	.38	.59	.101	.007
Number of Residences in One Year	.102	.26	.63	.105	.004
Prior Arrest for a Property Offense	.180	.58	.47	.109	.004
Prior Conviction on a Prop. Offense	-.097	.30	.66	.112	.003
Prior Arrest for Crime Against Person	.107	.15	.94	.114	.002
Age at First Arrest	-.167	.22	.76	.117	.003
Prior Sentence Including Probation	.141	.12	1.05	.119	.002
Substance Abuse Problem	.130	1.20	.09	.120	.001

* "Failure" is defined as rearrest, reconviction, or revocation during the first six months.

TABLE 22

ICFS RISK SCREENING INSTRUMENT:
INTERCORRELATION OF PREDICTORS

	1	2	3	4	5	6	7	8	9	10	11	12	13
1. Educational Attainment	1.0												
2. Age at Probation Sentence	.10	1.0											
3. Number of Residences	.05	.01	1.0										
4. Number of Prior Adult Arrests	.13	-.19	.17	1.0									
5. Number of Prior Sentences with Probation	.09	-.09	.08	.47	1.0								
6. Age at First Arrest	-.16	-.61	-.04	-.13	-.19	1.0							
7. Age at First Conviction	.17	.51	.06	.12	.23	-.76	1.0						
8. Substance Abuse	.05	-.09	-.03	.24	.17	-.11	.15	1.0					
9. Drug Abuse	.07	.11	-.02	.2	.15	-.22	.26	.62	1.0				
10. Current Charge	.00	.04	.05	.00	-.05	-.03	.03	.04	.01	1.0			
11. Prior Arrest for Offense Against Person	.09	-.11	.02	.28	.28	-.15	.13	.12	.05	.08	1.0		
12. Prior Arrest for Offense Against Prop.	.15	.04	.09	.40	.36	-.39	.38	.11	.16	.04	.22	1.0	
13. Prior Conviction for Off. Against Prop.	-.15	.03	-.02	-.33	-.35	.3	-.35	-.07	-.11	.01	-.21	-.78	1.0

The high correlations between variables using the age of the probationer at particular points may be expected in part from the circumstance that for nearly half the ICFS sample this was their first adult conviction. Other intercorrelations are rather modest.

As can be seen, the standard errors of the regression coefficients tend to be large relative to the coefficients, so it is questionable whether these weights may be used with confidence. Nevertheless, in view of our interest in using the resulting scores as a statistical control in the program evaluation analyses, on this same sample, the weights were retained. For the same reason, all 13 of these items were used.

The method of scoring for each probationer thus involved an equation consisting of thirteen weighted items. The weights used were the unstandardized regression coefficients. The items were those listed in Table 21.

Four of the variables were grouped into three categories (age at first arrest, age at first conviction, age at probation sentence, and educational attainment), while the remaining were included as dichotomous attributes. Each probationer thus received a risk score which was a weighted linear composite of the scores on these items. The scores provide an "expected" value of the "global" outcome criterion.

The distributions of these scores (called PREDRISK scores) are shown in Table 23 for the two outcome categories. It should be noted that this table shows the association of these scores with the global outcome for the samples from which they are derived, that is, the combined data from the three sites. The point biserial correlation coefficient of .35 is identical to the multiple correlation coefficient found; and, since they are equivalent statistics, this shows that there was no loss of predictive information in grouping the scores. Since we wish to examine the global outcomes for the three sites combined but also for each site separately, it is of interest to note the relation of these scores to outcome for each agency; these data are shown in Table 24.

For Tables 23 and 24, the PREDRISK scores were grouped to provide seven classifications of probationers. This was done merely for convenience in displaying and examining the association of the scores with the global outcome measure. These groups were defined arbitrarily by standard deviation units, such that the scores in group A are at or above two standard deviations above the mean score and B = 1σ to 2σ, C = $.5\sigma$ to 1σ, x = $-.5\sigma$ to $\sigma .5$, D = $-.5\sigma$ to -1σ, E = -1σ to -2σ and F $\leq -2.\sigma$ (σ is the symbol for the standard deviation).

TABLE 23

RELATION OF PREDRISK SCORES TO GLOBAL OUTCOMES,
THREE SITES COMBINED, STUDY SAMPLE

Group	Scores	Percent of Total	Number		Percent Favorable
			Favorable	Unfavorable	
A	1.087 or above	(<1)	1	0	100
B	108.6 - 96	14	66	4	94
C	95.9 - 88.6	21	94	10	90
X	88.5 - 75.5	37	166	20	89
D	75.4 - 68.8	12	43	18	70
E	68.7 - 55.6	13	39	28	58
F	55.5 - 0	3	8	9	47
Total		100	417	89	.82

Biserial correlation coefficient = .51

Point biserial correlation coefficient = .35

Mean cost rating = .46

Z score = 6.98

$p < .001$

TABLE 24

RELATION OF PREDRISK SCORES TO GLOBAL OUTCOMES FOR
SUFFOLK COUNTY, KANE COUNTY, AND FLORIDA

Site	Group	Percent of Total	Number		Percent Favorable
			Favorable	Unfavorable	
Suffolk	A	0	0	0	
	B	9	10	1	90
	C	16	19	1	95
	X	42	49	14	92
	D	17	18	3	86
	E	15	6	12	33
	F	3	2	2	50
Kane	A	0	0	0	
	B	2	2	0	100
	C	12	11	1	92
	X	35	29	7	81
	D	19	12	7	63
	E	26	18	9	67
	F	6	3	3	50
Florida	A	0	1	0	100
	B	21	54	3	95
	C	26	64	8	89
	X	35	88	9	91
	D	8	13	8	62
	E	8	15	7	68
	F	3	3	4	43

Site	Biserial Correlation	Point Biserial Correlation	Mean Cost Rating	Z Score	p (MCR)
Suffolk	.63	.43	.58	2.37	<.05
Kane	.35	.26	.31	4.49	<.05
Florida	.49	.32	.40	4.12	<.05

In Table 22, the percent of probationers in each score group who had favorable global outcomes is shown in the column at the extreme right, and it can be seen that the proportions of probationers who "succeeded" decreased with decreasing scores. The correlations of the PREDRISK scores with the global outcome classification are shown below the table, along with the mean cost ratings. In this case, the value of MCR = .46, with a companion Z score of 6.98 which is significant at the .001 level of confidence. Similar data are given for each site separately in Table 24. Each of these reported MCR's has an associated Z score that is significant at the .05 level of confidence. Together, these tables show that the PREDRISK scores in this sample are related moderately to substantially to the global outcome classifications, for each site separately and for the sites combined.

In Table 25, the PREDRISK scores have been grouped to provide three classifications of probationers (as used in the projects). These groups are again defined rather arbitrarily by standard deviation units. The scores in group A are one-half a standard deviation or more above the mean score, B = $-.5\sigma$ to $.5\sigma$, and C $\leq$ $.5\sigma$. Because of some loss of information, the MCRs are slightly lower; but, the Z scores are all significant

TABLE 25

RELATION OF PREDRISK GROUPS TO ICFS OUTCOME FOR ALL SITES COMBINED, SUFFOLK COUNTY, KANE COUNTY, AND FLORIDA

Site	Group	Percent of Total	Favorable	Unfavorable	Percent Favorable
All Sites	A	35	161	14	92
	B	37	166	20	89
	C	20	90	55	62
	Total		417	89	82
Kane County	A	14	13	1	93
	B	35	29	7	81
	C	51	33	19	63
	Total		75	27	.73
Suffolk County	A	24	29	2	94
	B	42	49	4	92
	C	34	26	17	60
	Total		104	23	.82
Florida	A	47	119	11	92
	B	35	88	9	91
	C	18	31	19	62
	Total		238	39	.86

Site	Mean Cost Rating	Z Score	p
All Sites	.43	6.64	$<.05$
Kane County	.29	2.27	$<.05$
Suffolk County	.50	3.82	$<.05$
Florida	.37	3.91	$<.05$

at the .05 level of confidence.

Conclusions

The risk screening instruments used in three sites for the Improved Correctional Field Services project all had serious methodological flaws. Nevertheless, there is evidence for some validity of these instruments; but that evidence varies among the sites.

Neither Kane nor Suffolk Counties developed their instruments on the basis of a study of samples that could be assumed to be representative of the ICFS target population. In Florida, this assumption might have been tenable; but possibly the construction sample was biased by overrepresentation of probationers from urban areas.

None of the sites reported reliability studies of the items included in the instruments, of scores based on them, or of the resulting probationer classifications. (In Suffolk County there was attention to the reliability of classification using the instrument as compared with probation officer judgments.)

Neither Kane County nor Florida conducted any study of the validity of the instrument before putting it to operational use in the project. Suffolk County completed a validation study, but that study was not limited to the ICFS target population. Thus, on initiating the ICFS project Kane County and Florida had only weak evidence as to the validity of the risk classifications that

provided the basis for the study; Suffolk County had evidence of substantial validity of the device for their probationers in general.

The ICFS project data can provide only a weak test of the validity of the instruments used, particularly in view of:

-A possible confouding risk and treatment;

-Small numbers of probationers available for study; and

-A short follow-up period.

The results of this study, which should be considered in the context of these limitations, suggest that:

-The instruments used by all three sites to classify probationers into three risk groups for differential supervision are of questionable utility for that purpose.

This does not mean that the instruments are necessarily wholly lacking in validity. The best estimates of validity, ignoring significance tests, are given by the mean cost ratios obtained, as follows:

Site	Number	Mean Cost Ratio	
		Instrument	Classification
Kane	102	.26	.16
Florida	269	.20	.15
Suffolk	124	.20	.10

One may reasonably speculate that the values of these coefficients might improve with a longer follow-up study period and the expected change in the base rate as more "failures" occur.

Coefficients with higher values were found for a "risk screening" device developed on the basis of the ICFS subjects, for subjects from the three sites combined and for each site separately. Claims that this device has greater validity however, would be quite unwarranted; it has not yet been tested on any validation sample.

Further validation study of the instruments in use in the ICFS project - with larger samples and with longer follow up are needed for a more definitive assessment. The PREDRISK instrument may, on further validation study, be found to provide some bases for improvement of the predictive efficiency of the various instruments being used.

NOTES

[1]This report is one of a series addressing various aspects of the evaluation of the Improved Correctional Field Services Project. Others deal with the development of the project, with issues of implementation of the program plans, with the effects of the classification/differential supervision design, with problems of assessment of probationer needs, perceptions, and preferences, and with the scaling of offense seriousness.

[2]More specific information on the research plan employed is given in companion reports. See Finckenauer, J.O. and Gottfredson, D.M., The Improved Correctional Field Services Project: A Case Study, April, 1981 and Gottfredson, D.M., Finckenauer, J.O. and Taxman, Faye, Risk, Supervision, and Recidivism: The First Six Months of Recorded Experience in the Improved Correctional Field Services Project, July, 1980 (unpublished report drafts in this series of reports).

[3]Studies of risk assessment devices in probation settings, relative to those in parole, are fairly rare. See Albanese, J., "Predicting Probation Outcomes: An Assessment of Critical Issues," in Gottfredson, D.M., Finckenauer, J.O. and Rauh, C., Probation on Trial, Newark, New Jersey: Rutgers University School of Criminal Justice, 1977, pp. 129-178. For reviews of the problems of criminological prediction more generally, see Mannheim, H. and Wilkins, L.T., Prediction Methods in Relation to Borstal Training, London: Her Majesty's Stationery Office, 1955; Gottfredson, D.M.,

"Assessment and Prediction Methods in Crime and Delinquency," in Task Force Report: Juvenile Delinquency and Youth Crime, Presidents' Commission on Law Enforcement and Administration of Justice, Washington, D.C.: U.S. Government Printing Office, 1967; Simon, F.H., Prediction Methods in Criminology, London: Her Majesty's Stationery Office, 1971. For a comparison of methods commonly used, see Gottfredson, S.D. and Gottfredson, D.M., Classification for Risk: A Comparison of Methods, U.S. Government Printing Office, in press.

[4]Gottfredson, M.R. and Gottfredson, D.M., Decisionmaking in Criminal Justice: Toward the Rational Exercise of Discretion, Cambridge: Ballinger, 1980, p.4. The discussion of these concepts is drawn from this book, pp. 5-9.

[5]Wisconsin Division of Correction, The Wisconsin Case Classification/Staff Deployment Project: A Two Year Follow-Up Report, Project Report 14, Madison: Wisconsin Division of Correction, July, 1979.

[6]For an extended discussion, see Gottfredson, M.R. and Gottfredson, D.M., note 4, supra, at pp. 5-9.

[7]This section is adopted from Gottfredson, D.M., note 3, supra.

[8]The works cited at note 3, supra, contain extensive references to discussions of these issues.

[9]See Cronback, L.J. and Gleser, G.C., Psychological Tests and Personnel Decisions, Urbana: University of Illinois Press, 1957

[10]Michel, P.E. and Rosen, A., "Antecedent Probability and the Efficiency of Psychometric Signs, Patterns, or Cutting Scores, Psychological Bulletin, 52, 1955, pp. 194-216.

[11]Whether or not a prediction method is "useful" depends also on the purposes of the classification application. The concept of utility in relation to selection and placement decisions is discussed extensively in Cronback, L.J. and Gleser, G.C., note 9, supra.

[12]Duncan, O.D., Ohlin, L., Reiss, A.J., and Stanton, H.P., "Formal Devices for Making Selection Decisions, American Journal of Sociology, 58, 1953, pp. 573-584.

[13]Lancucki, L., and Tarling, R., "The Relationship between Mean Cost Rating and Kendall's Rank Correlation Coefficient," in Gottfredson, D.M., Wilkins, L.T., and Hoffman, P.B., Guidelines for Parole and Sentencing: A Policy Control Method, Lexington: D.C. Heath, 1978, pp. 199-206.

[14]Note 3, supra; see especially Gottfredson, S.D. and Gottfredson, D.M., in press.

[15]Burgess, E.W., in Bruce, A.A., Burgess, E.W., and Harno, A.J., "The Working of the Indeterminate Sentence Law in the Parole System in Illinois," Springfield: Illinois Parole Board, 1928.

[16]When the outcome criterion is a qualitative variate

with two categories such as probation "success" vs. "failure", these are equivalent procedures. See, e.g., Bechtoldt, H.P., "Selection," in Stevens,S.S., Handbook of Experimental Psychology, New York: Wiley, 1951, pp. 1237-1266 and, for a demonstration of the relationships, Porebski, O.R., "On the Interrelated Nature of the Multivariate Statistics Used in Discriminatory Analysis," British Journal of Mathematical and Statistical Psychology, 19, 2, 1966, pp. 197-214, p. 202.

[17]See Gottfredson, S.D. and Gottfredson, D.M., and Simon, F., note 3, supra; and Wainer, H., "Estimating Coefficients in Linear Models: It Don't Make No Nevermind", Psychological Bulletin, 83, 1976, pp. 213-217. A summary of the Gottfredson and Gottfredson report in press is given under the same title in Gottfredson, S.D. and Gottfredson, Criminal Justice and Behavior, 7, 3, September 1980, pp. 315-330.

[18]Gottfredson, S.D. and Gottfredson, D.M., note 17, supra.

[19]Ahasic, A.R., "A Risk Assessment Instrument for 16th Judicial Circuit Court Services, Kane County, Illinois," an unpublished paper.

[20]See Lancuski, L., and Tarling, R., note 13, supra.

[21]Gottfredson, D.M., Finckenauer, H.O., and Taxman, F., note 2, supra.

[22]Golbein, J.J., "Differential Classification for the Supervision of Adult Probationers: An Operational Design," Probation and Parole, 1977.

[23]Bonn, J.P., "Quarterly Report, Improved Correctional Field Services," Albany: New York State Division of Probation, July 16, 1979.

[24]Golbein, J.J., personal communication, April 4, 1980.

[25]Golbein, J.J., note 22, supra, at p. 191.

[26]Bonn, J.P. note 23, supra.

[27]See, for example, Pritchard, D.A., Stable Predictors of Recidivism, Washington, D.C.: American Psychological Association, and the reviews cited at note 3, supra.

END