THE DIAGNOSTIC PAROLE PREDICTION INDEX

BY

ERNST WENK

DECEMBER 1979
RESPONSIBLE ACTION, INC.
P. O. BOX 924
DAVIS, CALIFORNIA 95616

This research was funded under Order No. 8-0235-J-LEAA from the National Institute of Law Enforcement and Criminal Justice, Law Enforcement Assistance Administration, United States Department of Justice, to Responsible Action, Inc.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

I. INTRODUCTION

The Diagnostic Parole Prediction Index (DPPI) project was undertaken under LEAA Grant Award Number 75-NI-99-0039 to determine whether a parole prediction instrument combining statistical prediction with clinical case-study concerns could be developed using a clinical-synthesis model. The primary goals of this research were (1) to demonstrate that predictive information regarding parole outcome can be fused with clinical concerns in a device relevant to decision making within the case-study process; (2) to compare the accuracy of several prediction methods (multiple regression, predictive attribute analysis, association analysis, and the Burgess method) in forecasting parole outcome; and (3) to design and test in the field several alternate formats for figuratively displaying DPPI information in a manner that is readily understandable and useful to case workers.

Utilizing an existing data base consisting of information on variables collected on 4,146 California Youth Authority (C.Y.A.) wards (with fifteen-month post-release follow-up data), this project involved the following tasks: (1) the derivation of seven information dimensions or variable categories believed relevant to the clinical context (individual case history, offense-specific, academic, intelligence, vocational, social/psychological, and

psychological areas); (2) the comparison of four prediction strategies in terms of their appropriateness to the clinical format of the DPPI, as well as their predictive efficiency; (3) the design of three alternate formats for presentation of dimensional data; and (4) the field-testing of the three formats to determine their relative utility and appropriateness for decision making by the clinical practitioner.

It was expected that a "trade-off" between predictive efficiency and clinical utility would have to be made in order to develop an instrument that was both acceptable to the field and accurate in its statistical predictions of parole outcome.

The model developed was not intended to be a definitive instrument which would provide an inflexible synthesis of concerns derived from the project's rationale. Although the concerns which would determine the form of the final model are common to any state or jurisdiction, they vary in detail according to the goals of the specific agency and to the perspectives of the correctional personnel in that agency. This study used correctional staff of the California Youth Authority as the primary source of feedback on three prediction method formats (Burgess, multiple regression, and predictive attribute analysis). No attempt was made to ensure that the people involved in the field evaluation were a random sample of correctional personnel.

This project was interested in individual decisions, the process by which case workers arrive at them, and the methods by which a model predictive device might affect the process to a greater extent than previous prediction techniques have done. The DPPI thus is not a specific instrument but rather a process by which prediction and interpretation can be further integrated. Knowing the pitfalls of attempting to achieve explanation from correlation, the researchers have made a cautious attempt to derive further "meaning" from an actuarial device. To derive this meaning, the model was designed to provide predictive information in a way which would confront the decision-making process on two levels by: (1) allowing the identification of "leads" for more intensive interviewing, and (2) facilitating self-interrogation on the relationships between different groups of predictive information and the criterion of parole outcome. The decision maker could thus identify information elements felt to be important in determining parole success or failure and could also assess subjective hunches.

In this study, "diagnosis" refers to the process of identifying other interviewing variables which may be important to consider in the further investigation of a specific case. Although prediction does not necessarily provide an explanation of why a person acts in a certain way, prediction information can provide the impetus for attempting

to ascertain the causes of behavior and the basis for the construction of further questioning during an interview. Assuming that parole outcome is the criterion of concern to correctional staff, the decisions which affect this criterion are of great importance.

II. THE STATE OF THE ART IN PAROLE PREDICTION

A. Historical Perspective

For more than fifty years, researchers have been studying the prediction of criminal behavior. Numerous efforts have been made to estimate accurately the degrees to which different persons are at risk of committing offenses. The term "prediction" is used to mean any kind of estimation of the probability of the future occurrence of one event from a knowledge of factors to which it is related. While prediction studies in criminology have tended to construct and utilize predictive instruments derived by actuarial, or statistical, methods, it is important to remember that non-statistical methods do exist (e.g., clinical estimations of probability by correction workers or other justice system personnel) and that they are more frequently used in the field than are actuarial prediction tables. One of the fundamental issues in the field of prediction is that of the relative merits of "clinical" and "actuarial" prediction methods.

Definitions of prediction in criminology usually refer to actuarial prediction. For example, Lejins (1962) defined parole prediction as "the estimate of probability of violation or nonviolation of parole by an offender on the basis of experience tables, developed with regard to groups of offenders possessing similar characteristics."

Simon (1971) defined the prediction instrument as "one which uses certain information applying to a person at one time in order to estimate the probability of his becoming, or remaining, criminal (or delinquent) at some later time." The criterion of criminality must be clearly defined. She stated that the prediction instrument must distinguish between different risks; that it uses the information on which it is based to classify persons according to their different probabilities of becoming criminal; and that it must separate the low risks from the high ones. The instrument may take various forms: a table, a score, or an equation giving individual probabilities of risk. Prediction instruments are constructed for defined classes of persons or specific populations and may be based on information of several kinds: psychological or other test scores, biographical data, case material generated by the direct observations of clinical staff, etc. The selection and combination of predictive items may be accomplished by a variety of mathematical techniques, ranging from a simple scoring of points

for favorable and unfavorable factors to more complex methods which take into account not only the association of each factor with subsequent criminality but also the relationships between the factors themselves.

The steps to be followed in prediction studies have been outlined by Gottfredson (1967): (1) the establishment of criterion categories (e.g., parole outcome); (2) the selection of predictor candidates; (3) the testing of the relationship between (1) and (2), which yields an "experience table"; (4) the cross-validation of the table on new samples, which leads to the creation of the prediction instrument; and (5) the application of the prediction instrument. In the United States prediction tables generally are called "base expectancy" tables because of the early development of such tables to provide a base for further research by quantifying expectations (Gottfredson and Beverly, 1962). Most frequently such tables are developed atheoretically and intentionally omit information about institutional programs and performance. Gottfredson and Beverly (1962) define the base expectancy as a statement of the expected parole success rate for a given group made on the basis of past experience with other similar groups.

B. Development of Parole Prediction

One of the first attempts to introduce an actuarial

prediction method into parole was the "experience table" of Hart (1923), who is considered to be the originator of the concept of parole prediction. In 1923 Warner concluded that life history and background factors were of little value in predicting parole outcome. Reanalyzing Warner's data, Hart found that the accuracy of the prognostic score could be significantly improved if individual predictors were pooled. While the techniques of prediction by experience tables have been improved since Hart's analysis, the method has remained essentially the same.

In 1928, Burgess undertook the first large-scale investigation of the relationship between offenders' background factors and parole outcome. His study resulted in the development of a prototype expectancy table, which was introduced into the Illinois parole system in 1933. Burgess' table was derived from computations of the degree to which violation rates of subpopulations with specific background characteristics deviated from the average violation rate of a given parolee population. Where the subpopulation violation rate was lower than that of the total parolee population, the corresponding background factor was considered a favorable one. All positive factors were incorporated into an experience table and a candidate for parole was assigned one point for each favorable factor in his background. A table giving the violation rate for

offenders with different numbers of favorable factors was derived for the population under study.

In 1930, the Gluecks introduced the idea of weighting background factors according to the degree of their relationship to parole outcome. These weights, derived through statistical techniques, are more precise than the +1 or 0 weighting employed by Burgess. However, modern work on test construction and item weighting has indicated that the level of efficiency is not appreciably higher for scales using complex weighting systems than for those using simpler weightings (Gough, 1962).

Vold (1931), examining the correlation of each of 34 pre-parole factors with parole violation, found that, while none was of outstanding importance, most had at least some predictive utility. To determine whether combining factors of relatively high or low individual associations with parole violation would produce more effective predictors, Vold compared two methods of combination: the Glueck (weighted) scoring method and the Burgess point system. He then worked out 27 tables which, compared to the results achieved in most other studies, showed remarkably high predictive power (Simon, 1971).

Laune (1936) observed that since Burgess-type prediction was based on pre-institutional factors (such as work history, marital status, etc.), which do not change in

response to treatment, the correctional process was hardly considered in predicting parole outcome. Laune suggested the introduction of dynamic factors that are subject to change during the offender's incarceration. However, Lejins (1962) noted that, while Laune's research is important, the follow-up of his study indicated that his approach was no more effective than the Burgess method.

Tibbetts (1931) and Sanders (1935), among others, noted that predictions derived from different parolee populations were not necessarily consistent and that changes in administrative policies or in the general conditions of inmate life could affect the role of background factors. Ohlin (1951) later pointed out the need to continuously adjust the experience tables through an ongoing incorporation of research on predictive factors.

The Gluecks' research on the prediction of juvenile delinquency (1950) led to the development of three types of prediction table: one based on social history factors, one incorporating aspects of character structure as rated on the Rorschach test, and the third based on psychiatric evaluation of traits of temperament. The Gluecks (1950) recommended the use of the social history table, which is much easier to apply and is about as valid as a combination of the three tables. This table was originally constructed from matched samples of 451 institutionalized delinquent

boys and 439 nondelinquent boys. Five of the factors which distinguished between delinquents and nondelinquents were selected for inclusion in the Social Factors Prediction Table. Each category of these variables was scored by attaching to it the delinquency rate, expressed as a percentage, of boys within the category, providing a "weighted failure score" for each boy. The scores were then grouped and group failure rates were calculated. This approach has been used by the Gluecks in most of their prediction work. However, many writers (e.g., Gough, 1962; Voss, 1963; Reiss, 1951) have noted that weaknesses in the construction and testing of the Social Factors Prediction Table limit its confident application.

The post-World War II period was characterized by attempts at methodological refinement (Schuessler, 1954). Ohlin and his associates made significant contributions to this effort. For example, their "index of predictive efficiency" was designed to measure the percentage change in prediction error resulting from the use of an experience table instead of the overall rate (Ohlin and Duncan, 1949).

Ohlin and Duncan (1949), reviewing the results of fifteen prediction studies including work by Burgess, Vold, Monachesi, the Gluecks, and an early version of Ohlin's own parole table, introduced Predictive Efficiency (P.E.) as an index for measuring predictivity and showed that, for all

of the studies reviewed, predictive power shrank on application to the validation sample. Although P.E. has since been replaced by the Mean Cost Rating (M.C.R.)[1], their finding that predictive power is likely to shrink on validation remains true. They suggested that prediction error derives from three sources: lack of association between predictive factors and outcome in the population, sampling fluctuations, and changes in associations over time (Ohlin and Duncan, 1949).

One of the most thorough prediction studies was undertaken by Ohlin (1951), using data on over 17,000 parolees from Illinois prisons between 1925 and 1945. After trying various systems of scores and weights for combining predictive factors, he found that a simple points score such as Burgess had developed worked as well as any. He found also that a 12-factor version of Burgess' 21-factor scale predicted as accurately as the original. Ohlin's twelve factors were not restricted to pre-sentence items but included some that reflected the prisoner's situation during his current sentence.

In 1952 Ohlin and Lawrence published the results of a parole prediction study which used subjective data from an earlier study by Laune (1936). Laune had obtained "hunch"

---

[1] See pages 29, 64-65 for a discussion of Mean Cost Rating.

estimates from four Illinois prison inmates concerning the chances of success on parole for 150 of their fellow inmates. A large number of attitudinal and objective factors were extracted from these estimates and combined into a questionnaire which was administered to a large sample of inmates. A Burgess prediction score based on objective factors also was compiled.

Ohlin and Lawrence compared the predictive accuracy of the four inmates' original "hunch" estimates with that of the Burgess score, finding that two of the inmates did a little better than the Burgess score and one did considerably worse. They then divided into two samples 823 of the men to whom the questionnaire had been administered. From the first sample they constructed thirteen prediction instruments by a points scoring method, using selections from Laune's questionnaire factors and, separately, objective factors from the Burgess score or Ohlin's later development of it. When all thirteen were tested on the second sample, the Burgess-type objective factors, while more discriminating than the subjective factors on the construction sample, were found to be less stable on validation.

During the late 1940's and the 1950's there was much experimentation with and further development of techniques for statistical prediction using multiple regression equations, the discriminant function, and configural or non-linear methods.

The regression equation, the earliest and perhaps the most widely used technique, may be used, with two variables (one independent and one dependent), to indicate how many units one variable increases for every one-unit increase in the other or, with more than two variables, to predict the value of one from a combination of the others. For the prediction of certain criteria, particularly those which are quantitative, normally distributed, and demonstrably correlated with other variables, the regression equation has produced excellent results (Gough, 1962). Sarbin (1941) recommended it as a model for the thought processes to be followed by the clinician in individual diagnosis and found (Sarbin, 1942) that, in predicting students' scholastic achievement, college counselors did use a form of reasoning similar to multiple regression. Others (e.g., Chein, 1945; Klein, 1948) objected to its use in clinical prediction because of its assumption of linearity--i.e., that for every increment in one variable there will be a proportionate incremental change in the other. Hoffman (1958, 1959) later proposed that the regression method be used to determine the degree to which the clinician departs from simple linear functions in combining predictive information.

The discriminant function, which is used when subjects are to be placed in one of two discrete and nongraded classes,

has not been used frequently in clinical psychology (although there have been notable exceptions, e.g., Webster, 1952). One reason for this is that when the criterion is dichotomous it can be shown that the discriminant function is a simple linear transformation of multiple regression.

Some of the most interesting developments have been made in the area of configural and non-linear methods of actuarial prediction such as predictive attribute analysis (MacNaughton-Smith, 1963), configural analysis (Glaser, 1964), and association analysis (Wilkins and MacNaughton-Smith, 1954). These approaches seek to account for trait patterns, configurations of data, and non-linear relationships between predictors and criteria. One trend in this direction is the construction of indices of profile similarity (e.g., Helmstadter, 1957; Mosel and Roberts, 1954; and Muldoon and Ray, 1958), which permit estimates of an individual's correspondence to a criterion classification (or to his own behavior at another time, to the behavior of another person, etc.) to be derived from a comparison of his present profile of test scores and that of the criterion. Muldoon and Ray (1958) undertook a study that included clinical estimates of congruence among profiles with estimates derived from six statistical methods.

In 1955, Mannheim and Wilkins published what is now

one of the classics among prediction studies. From the case records of a representative sample of youths entering borstal during 1946-1947, they obtained about sixty variables, from which they selected those that were sufficiently reliable and significantly correlated with the parole failure criterion of reconviction. Multiple regression resulted in an equation combining five pre-sentence variables, from which a score was derived and a five-class table constructed. Mannheim and Wilkins compared the predictive power of their table with that of material contained in institutional staff reports and found that their objectively based table was more discriminating than the subjective clinical assessments. Numerous others (e.g., Benson, 1959; Gibbens, 1965; Hood, 1966, Cockett, 1967) have used the Mannheim-Wilkins table, or some modification of it, in their research.

Using the theory of "differential identification," Glaser (1954, 1955) attempted to derive predictive factors from the parolee's previous identifications with persons of conventional or unconventional values and his economic opportunities and acceptance among conventional associates. Using a weighted scoring system, Glaser combined into a prediction table seven factors whose predictive utility was high and relatively consistent for parolees released during three time-periods.

Glaser (1962) presented a new method of combining factors, which he called a configuration table. This technique had been suggested by other writers and is similar to MacNaughton-Smith's (1965) predictive attribute analysis. Later experience with the configuration table has shown that it is subject to substantial shrinkage on validation (Simon, 1971).

In 1962, Gottfredson and Beverly summarized the work of preparing "base expectancy" tables for adult prisoners and Youth Authority wards in California. Such tables have been used extensively by state correctional agencies in California both in research and in program decisions. Gottfredson and Beverly presented three tables indicating the chances of success on parole for men and women within two years after release and for boys within 15 months. Each table was derived by multiple linear regression analysis, the regression equations being transformed into scores which were then grouped to form a table.

Gottfredson and his associates (1963) compared three techniques--multiple regression, association analysis, and predictive attribute analysis--by analyzing data on the same sample. On validation, this analysis produced tables of seven and ten classes, which included some of the variables from the regression equation. Gottfredson and Ballard (1965) tested the tables constructed by multiple

regression and association analysis on the same sample, finding that they predicted as well after eight years as after two.

Beverly (1964) compared the predictive value of data found in the routine records of all California Youth Authority wards with more extensive data on family background and personal history gathered by social workers during visits to the wards' homes. Using multiple regression on a construction sample of 3,046 cases, Beverly developed two tables, one from the routine data alone and the other with social history information added. He found that the improvement gained by incorporating the social history data was statistically significant.

Gough, Wenk, and Rozynko (1965) compared one of the CYA base expectancy tables with personality inventories. Using a construction sample of 444 CYA parolees and a validation sample of 295, they developed six multiple regression equations from selections and combinations of the variables contained in the boys' base expectancy scores, the Minnesota Multiphasic Personality Inventory (MMPI), and the California Psychological Inventory (CPI). These equations were compared with each other, with the base expectancy score alone, and with several earlier prediction scales.

When the discriminating power of the equations was assessed, the best results were obtained with the combination

of the base expectancy and the CPI. This equation, as well as two other combinations (the B.E. + MMPI and the B.E. + MMPI + CPI) did better than the B.E. alone. To measure their predictive accuracy, the equations were transformed into two-class tables, predicted successes and failures. The best equation, B.E. + CPI, produced a "hit" rate of 63 percent, compared to 59 percent for the B.E. alone. Gough and associates pointed out that one advantage of combining inventory data with the B.E. score was that while the B.E. score was based largely on previous criminal record determined at admission, inventories could be administered during institutional confinement and thus reflect changes which resulted from treatment and which might affect parole outcome.

Gottfredson (1967) outlined the nature and problems of prediction in crime and delinquency and noted that further research was needed into the empirical comparison of the predictive efficiency of various methods for combining predictors. Other suggested areas for research included: the incorporation of criminological theory into prediction work, resulting in contributions to both theory and prediction; the testing of hypotheses derived from clinical practice; the examination of the possibilities for collaboration, rather than competition, between statistical and clinical prediction methods; the improvement of criterion

measures; the cross-validation of existing measures; the development of measures of specific subgroups rather than samples of total populations; and the incorporation of prediction methods into the information systems of agencies responsible for offender management.

Babst, Gottfredson, and Ballard (1968) compared multiple linear regression and configuration analysis as techniques for constructing parole prediction tables. They found that the tables produced similar results with the same data, although the configuration table used only three variables while the regression table used six. Babst and his associates concluded that both methods worked about equally well and suggested that prediction might be improved by a combination of the two.

Ward (1968) compared five analytical techniques: a weighted points score, the Burgess (unweighted points score), the Glueck method, discriminant function analysis, and multiple linear regression. He found the least correlation for the Burgess, which is the simplest scaling method; those which weight factors but ignore intercorrelations were next; and scales which account for correlations between factors were the best.

In 1968, Dean reviewed the literature on parole prediction, noting several weaknesses in prediction research: (1) the variables used in prediction devices were subject

to the same criticisms leveled against those used in the Burgess system three decades earlier (i.e., they were static, extrinsic to the individual, and restricted to data collected in prison files for administrative purposes); (2) a point of diminishing returns had been reached in applying more rigorous analytical techniques to such data; and (3) with a few exceptions, there had been no effort to relate this research to the rather substantial body of criminological theory which had been developed. With a view toward rectifying this situation, Dean presented some hypotheses, suggested by various widely accepted criminological theories, as possible new directions for prediction research.

Building on the theoretical insights of Glaser (1960), Rogers (1967) examined the degree of congruence of sociological theory, general research results, and the perceptual accuracy of correction workers with respect to certain prediction items. Noting that correctional practitioners had not responded favorably to parole prediction devices designed by researchers for their use, Rogers attempted to determine whether the ways in which correctional personnel viewed the correlates of parole success corresponded to the findings of independent research on these correlates. From this perspective, it was suggested that correctional decision makers might reach conclusions about parolees that

coincide with the findings of research and act accordingly without making formal use of prediction tables. Rogers gathered data on 20 parole prediction items from over 400 respondents representing ten professional positions at nearly all points of the classical correctional sequence. The items were derived from various theoretical rationales, such as differential association, differential alienation, differential anticipation, and norms from within the legalistic tradition. From this exploratory study, Rogers concluded that while there had been a lag in the formal use of prediction instruments, there was demonstrable congruence among theory, research findings, and perception of certain parole prediction items, especially those derived from differential association theory. Overall, more items were correctly perceived by correctional practitioners than were incorrectly perceived. In fact, those items which theory and research had found to be unfavorable to parole success were accurately perceived by correctional personnel. Rogers concluded that theory, research, and perceptions are all relevant to parole prediction studies.

C. Uses of Prediction Instruments

Prediction instruments are useful in research (e.g., treatment program evaluation and comparisons) and in administrative decision making (e.g., selection for parole

or assignment of offenders to appropriate programs). In correctional research a primary use for prediction instruments is in the evaluation of treatment programs. Offenders are classified into categories according to their likelihood (based on pre-treatment characteristics) of violating parole. This estimate of risk prior to treatment is then used as a baseline against which the results of treatment can be measured: programs are "effective" to the extent that they reduce the violation rate below the expected risk.

There are numerous problems, however, associated with the use of prediction instruments in treatment evaluation (Wilkins, 1969). These problems may be partly responsible for the relatively limited use of prediction tables in correctional systems. One of the exceptions has been California, where the development and use of prediction tables has played an important role in correctional research. The Community Treatment Project of the California Youth Authority, for example, involved the use of base expectancies in the comparison of direct release and special supervision in the community with the regular institutional program (Gottfredson and Beverly, 1962). Using the base expectancies to allow for the various risks, Beverly and Guttman (1962) compared, in terms of parole violation rates, the various institutions from which California Youth Authority wards were released.

In theory, prediction tables could be used by correctional decision makers and administrators at various stages: prevention, sentencing, treatment, release, and aftercare. In practice, although numerous instruments have been constructed by researchers, their use by practitioners has been limited. Evjen (1962) found that forty-four of the forty-eight states responding to an inquiry indicated that they never had used prediction instruments in parole selection and were not then using them. There is little indication that this situation has changed much in the intervening years (Simon, 1971).

There are several reasons for the reluctance of correctional practitioners to endorse the wide use of actuarial prediction instruments. A primary one is that the predictive power of these instruments has generally been rather low. Many correctional decision makers seem to feel that an experienced practitioner can make more accurate prognostications about an offender than can be derived from statistical tables (despite repeated demonstration, discussed in the next chapter, that this is not true) and that the use of prediction tables or risk categories is antithetical to the widely accepted goal of individualized treatment (Powers, 1962).

III. RATIONALE FOR THE DEVELOPMENT OF THE DIAGNOSTIC PAROLE PREDICTION INDEX

A. Background

Although previous research has achieved some success in predicting such criteria as parole outcome, it has not addressed any substantive issue except whether the clinician might improve upon the actuarial prediction. This kind of question can be and has been answered empirically, with the evidence indicating that the case worker cannot improve upon the statistical prediction. While some researchers have established the need to determine other ways these concerns might supplement each other, few practicable suggestions have been offered.

From its inception, the DPPI project has assumed there are previously untested methods of fusing actuarial or statistical prediction and case-study concerns. This chapter attempts to provide both the basis and the rationale for a model for combining actuarial and clinical concerns. The model is assumed to be theoretical enough to be applicable in correctional agencies other than the California Youth Authority (CYA), the site of the present study.

The assumptions underlying the model involve a set of hypothetical suppositions which, although plausible, are essentially untested. In fact, there is an intertwining of concerns, which in total form the rationale of the model. Figure 1 provides a flow chart of these concerns and their

interrelationships and shows the three levels of the rationale and their relationships to the various concerns.

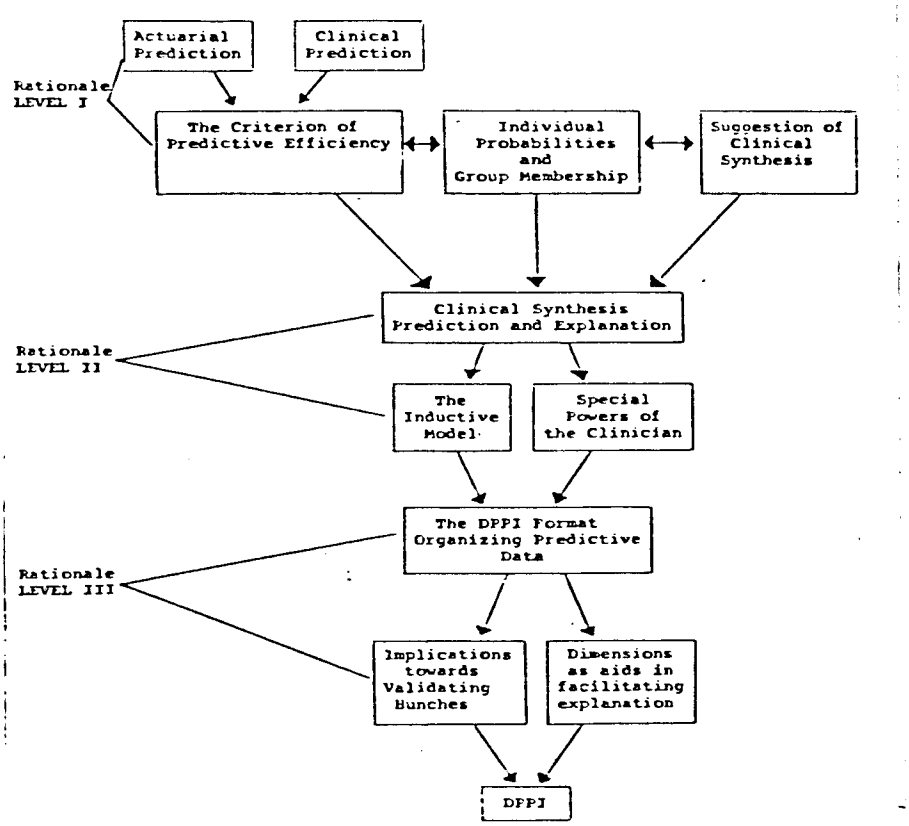


FIGURE 1
DPPI Rationale and Substantive Framework

The first level of rationale deals with the reasons for the possible usefulness to the clinician of actuarial prediction. The second level of rationale includes explanations of how the clinician might build upon predictive information. The third level of rationale deals with the instrument itself and with the reasons for its style

of construction as they follow from perceived limitations of previous predictive instruments. The substantive concerns and the sequence of their presentation (from top to bottom) provide an overview of the relationships among the concerns which form the instrument.

B. Level I Rationale: Actuarial and Clinical Prediction

i. Comparing Predictive Efficiency

Early in the development of the study, it became obvious that previous research had made little attempt to fuse the concerns of the two sides in the actuarial vs. clinical prediction controversy. While the two factions had been competing for forty years to predict various types of criteria, little time had been spent in understanding how these concerns might supplement each other.

The debate as to whether the actuary or the clinician is more accurate in predicting different outcomes has at various times been one of the major methodological issues of psychology. Although we may question whether these two approaches to prediction are fundamentally different, there are some obvious superficial distinctions between them.

Meehl (1954) provided a definition of the actuarial or statistical approach to the predictive task:

> We may order the individual to a class or set of classes on the basis of objective facts concerning

his life history, his scores on psychometric tests, behavior ratings or check lists, or subjective judgments gained from interviews. The combination of all these data enables us to classify the subject; and once having made such a classification, we enter a statistical or actuarial table which gives the statistical frequencies of behaviors of various sorts for persons belonging to the class. The mechanical combining of information for classification purposes, and the resultant probability figure which is an empirically determined relative frequency, are the characteristics that define the actuarial or statistical type of prediction. Meehl then elaborated on the clinical approach: Alternatively, we may proceed on what seems, at least, to be a very different path. On the basis of interview impressions, other data from the history, and possibly also psychometric information of the same type as in the first sort of prediction, we formulate, as in a psychiatric staff conference, some psychological hypothesis regarding the structure and the dynamics of this particular individual. On the basis of this hypothesis

> and certain reasonable expectations as to the course of outer events, we arrive at a prediction of what is going to happen. This type of procedure has been loosely called the clinical or case-study method of prediction.

The actuarial approach has at different times been referred to as mechanical, objective, and quantitative while the clinical approach has been described as non-mechanical, subjective, and qualitative. Gough (1962) claimed that the difference between the two is operational: "...If the procedures, however complex mathematically, are such that anyone could carry out the necessary operations ...the method is statistical or actuarial.... If the combining is done intuitively...and constructs are generated during analysis...by judgment and reflection, the method is clinical." Sawyer (1966), too, noted that "The distinction between the clinical and actuarial method of prediction is found in the way in which the data, once specified, are combined for use in making the prediction."

There are several techniques for comparing the predictive efficiency of clinical and statistical approaches to prediction. The most commonly used method is a "coefficient of agreement," in which actual outcomes for each individual are correlated with the predicted outcomes. The higher the resulting coefficient, the more accurate the method. Several

studies have used a somewhat cruder measure, a simple "hit" percentage indicating the number of correct predictions out of the total number of predictions made.

Other, more complex, measures of predictive efficiency have been developed from probability theory, notably by Ohlin and Duncan (1949), Duncan et al. (1953), and Duncan and Duncan (1955). To be useful, a prediction table or method must be able to predict better than the base rate prediction for an entire sample. If, for example, 30 percent of all parolees become violators, then a "prediction" that all parolees would be non-violators would be correct 70 percent of the time. A useful predictive device therefore would have to be correct more than 70 percent of the time. Duncan and Duncan (1955) developed the Mean Cost Rating (MCR), a technique designed to reflect the degree to which a classification method succeeds in differentiating between those who risk being violators and those who do not (or between the two parts of any such dichotomous criterion).[2]

---

[2]The MCR is one of the techniques used in this study to compare actuarial methods of prediction. The MCR is at a maximum of 1.0 if all cases are accurately classified into two groups, one with 100 percent successes and the other with 100 percent failures. It is at a minimum of zero when all categories of cases distinguished have the same success or failure rates.

Research has tended to show that efficiency of actuarial predictions is somewhat greater than that of clinical predictions. Burgess (1928) found that an experience table was somewhat more accurate than probation officers' predictions of success on probation. Wittman (1941), Sarbin (1942, 1944), and Wittman and Steinberg (1944) also found the statistician to be more accurate in predicting various criteria. Mannheim and Wilkins (1955) and Thompson (1952) too obtained results which favored the actuarial method. Two somewhat more conclusive studies were undertaken by Meehl (1954) and Sawyer (1966). Meehl found that "Of 27 studies, 17 showed the statistical method clearly superior; ten showed the methods of equal efficiency." In a comparison of 45 studies, Sawyer found the statistical method of combining data to be superior to the clinical method in predicting various behavioral outcomes.

Many clinicians have claimed that comparisons of actuarial and clinical prediction are inherently unfair to the clinician. McArthur (1968) stated that "clinical prediction" is a misnomer, since the primary goal of the clinician is not to predict specific outcomes of persons, but rather to understand human nature. He remarked:

> Phrased more abstractly, our questions were how personalities are structured and why and how we

could hope to discern the laws governing the behavior of any given individual. These are nontrivial questions. We need to know many human natures before we can hope that Human Nature will be revealed to us.

These questions are "clinical." They seek to understand the dynamics of persons--not the correlations among institutionalized events.

Polanyi (1964) noted that prediction in itself is an insufficient basis for understanding, since "...correlation does not necessarily imply causation." Holt (1970) observed that "...the logic of statistical prediction does not require understanding of behavior...only the correlation between the predictor and criterion...no wonder statistical prediction has made such a small contribution to psychology."

The preceding discussion suggests differences in the assumptions underlying the clinical and actuarial approaches, as well as differences in the processes by which each arrives at a prediction. It will be seen that the difference in process is a key element of the suggested fusion. At present, however, it is sufficient to note that the clinician's prediction is a by-product of an extensive effort to understand the individual.

ii. Combining Actuarial Prediction and Case Study Concerns

Thus far the desirability of using actuarial techniques in the case-study context has been indicated. However, the criterion of predictive efficiency should not be the only criterion to be considered important. Parole practitioners often have complained that predictive devices, while identifying groups of individuals having a certain probability of parole success, do not consider an individual's probability of parole success. This criticism can be assessed by looking at the nature of probability.

Allport (1940) stressed the difference between predictions for a population and for a single case:

> Suppose we set out to discover the chances of John Brown to make good on parole, and use for the purpose an index of prediction based upon parole violations and parole successes of men with similar histories. We find that 72% of the men with John's antecedents make good, and many of us conclude that John, therefore, has a 72% chance of making good. There is an obvious error here. The fact that 72% of the men having the same antecedent record as John will make good is merely an actuarial statement. It

tells us nothing about John. If we knew John sufficiently well, we might say not that he had a 72% chance of making good, but that he, as an individual, was almost certain to succeed or else to fail.

In this statement Allport made two important assumptions: (1) that prediction for a group and prediction for an individual are separate and exclusive tasks, and (2) that determining the subclass with an associated parole success rate tells us nothing about John Brown's probable success on parole.

Table 1 provides the relative parole success figures, derived actuarially, for five groups of offenders.

Table 1
Hypothetical Parole Success Rates for Five Groups of Offenders

Subgroup	Number		Percent	
	Success	Failure	Success	Failure
A	16	4	80%	20%
B	12	8	60%	40%
C	10	10	50%	50%
D	8	12	40%	60%
E	4	16	20%	80%

Let us assume that John Brown is a member of subgroup A,

which has an overall parole success rate of 80 percent. According to Allport's argument, even though we know the overall success rate for that group, we know nothing about John's probable parole outcome. Allport implied that if we knew John sufficiently well we could predict his probable parole outcome by non-actuarial methods. In this way, we might determine that his probability of success on parole is higher than 80 percent. The important point here is that, even if we knew John well, we would most likely consider his success on parole to be less than certain--i.e., the outcome would still be a matter of probability. We are therefore using an implicitly actuarial notion.

Sarbin (1944) maintained that predictions about a single case in clinical work are never certain, but are always probable. Statements about the probability of a given event are statements about frequencies (95 percent certain means 95 chances out of 100). Frequencies refer to the occurrence of events in a class; thus, all predictions, even those that are predictions about individual persons, refer to a class.

The assumption that some form of probabilistic notion underlies the prognostic statements of the clinician has led some authors to suggest that clinicians seek to make their "...predictions in terms of personal probabilities"

(Pankoff and Roberts, 1968). Winkler (1967) indicated that self-interrogation could reduce the vagueness of subjective probability, allowing a better awareness of the process of applying probabilities to individuals or events. The primary point here is that, in stating that he is "almost sure" that John Brown will succeed on parole, the case worker should understand that he is making a probabilistic statement which could be interpreted in terms of a percentage.

In principle, all laws, even of the causal-dynamic type, refer to classes of events (Meehl, 1954). Paradoxically, the uniqueness of individual events forces one to assume that it is rational to entertain future expectancies on the basis of class membership, since the alternative is to conclude that nothing can be said about John Brown on the basis of class membership. In terms of probability, John's association with a group which has an 80 percent probability of parole success does tell us something about John since, given no other information about him, we can conclude that he also has an 80 percent chance of succeeding on parole. We suspect that even Allport would not deny the rationality of predicting the individual subject's behavior if a regression system led to a multiple R of .999 between a group of characteristics and parole outcome. With a

perfect R of 1.00, could we hesitate to apply the prediction to an individual in this class? If this is reasonable, are not .990 or .90 and thus .75, etc. also reasonable?

Given that the actuarial table does tell us something about John Brown's probability of parole success, and given that prior research indicates the superiority of the actuarial table in predicting specific outcomes such as parole success, why can we not supplant the clinician or case worker with such a device? To answer this question it is necessary to look closely at the clinical process. In accordance with the assumptions of clinical synthesis, as defined by Sawyer (1966)[3], a major assumption of the D.P.P.I. project is that the case worker should be able to build upon the classification of John Brown as someone with an associated 80 percent probability of parole success. We shall now examine the nature of the clinical element in the fusion process.

The process of determining John's probability of parole success involves a sequence of steps whereby we seek to understand John in greater detail. Even if we were dealing with a totally subjective procedure of determining whether John will succeed or fail on parole, we would feel

---

[3]See pages 38-40 for a discussion of Sawyer's work.

confident only after we knew a certain amount about John. It is obvious therefore that while a predictive instrument may give us some measure of John's likelihood of succeeding on parole, it cannot provide us with information that indicates John's specific chances of success on parole.

Previous research has suggested that actuarial and case-study concerns might supplement each other. Holt (1958), Hutt (1956), and Zubin (1956) suggested that the clinician can formulate relationships, but that in so doing he should be guided by actuarial frequencies and statistical analyses. Coyle (1956) and Trankell (1959) observed that "...even though the actuarial system excels the clinician in its general baseline accuracy rate... this rate might be augmented by adding clinical judgment as a separate factor." Other researchers, such as DeGroot (1960), have claimed that there is more to be gained by attempting to develop a basis for participation between the two concerns than by promoting further competition. Even Allport (1961) envisioned the clinician as being able to integrate an individual's characteristics into the actuarial scheme, while Sawyer (1966) stated, "There is after all no inherent reason to withhold from the clinician any relevant information - even the actuarial prediction itself...." Finally, Gough (1962) noted:

"Proper use of the clinician's skills might well be as a supplement or addition to the forecasts of the regression equation...the superior accuracy of the prediction equation might be augmented by adding clinical judgement as a separate factor."

Although there has been interest in combining actuarial and clinical concerns, few attempts have been made to do so. Meehl and Dahlstrom (1960) and Klopfer et al. (1951) attempted to derive complex configural judgments which could be incorporated into an actuarial scheme. Sawyer (1966), comparing clinical and actuarial accuracy in predicting different criteria, noted that the process of prediction involves two stages: the collection of the data and the combining of the data to make a prediction. At each stage the clinician has a unique role to play. The interview allows him to probe, to follow up cues, and to tailor his examination to the individual being studied; he can thus collect data which are not available from biographical records and tests. Collection of the latter, which is done according to specified rules and without involving clinical judgment, Sawyer calls "mechanical." At the second stage, data can be combined by mechanical rules to give an actuarial prediction, or they can be integrated into a clinical judgment.

From the two stages, Sawyer derived eight methods of collecting and combining information to arrive at a prediction: (1) pure clinical--clinically collected and clinically combined data; (2) trait ratings--clinically collected data which are mechanically combined; (3) profile interpretation--mechanically collected data, clinically combined; (4) pure statistical--mechanically collected data, mechanically combined; (5) clinical composite--both kinds of data, clinically combined; (6) mechanical composite--both kinds of data, mechanically combined; (7) clinical synthesis--taking a prediction produced mechanically and treating it as a datum to be combined clinically with other data; (8) mechanical synthesis--taking a prediction produced clinically and treating it as a datum to be combined mechanically with other data.

C. Level II Rationale: The Clinical Synthesis Model - Basis for the DPPI

Of the eight methods defined by Sawyer (1966), the seventh or clinical synthesis method is closest to one of the major goals of this study, i.e., the construction of a predictive device which the case worker can combine with clinical data. In comparing the predictive efficiency of the clinical synthesis method with the seven other methods, Sawyer found that the clinical synthesis approach did better than the clinical composite, although it did not do as well

as mechanical synthesis, in which the prediction of the clinician is added as a variable to be combined mechanically in the predictive equation.

Sawyer's study found that the addition of clinical judgment to the mechanical prediction did not improve the efficiency of the predictions, at least in the few studies where comparisons were possible. Gottfredson and Beverly (1962) also found that the subjective ratings of clinicians added nothing to the predictive efficiency of a simple checklist. Although these two studies cannot be considered conclusive, they indicate that the clinical synthesis model may not provide much more predictive accuracy than the actuarial instrument alone. Why, then, persist with clinical synthesis at all? It must be questioned whether the sole reason for the construction of the DPPI model is to increase predictive efficiency. To answer this question we must examine both the relationship between prediction and explanation and the interview process.

There can be little disagreement with the assumption that a primary goal of science is empirically testable explanation. It is true also that an important objective of the case-study or clinical process is to derive an understanding of an individual which allows the prediction of his future behavior. Of primary interest here are the

steps involved in the formulation of a clinical prediction. If the clinician's role were not highly creative, it would contribute nothing that a predictive device could not provide. However, the clinician is constantly forming hypotheses about individuals. There is a readiness to invent stimulus-response chains, which imply reference to theories. The clinical process thus involves the construction of "special theories" applicable to one person or to a few similar persons.

Sarbin and Taft (1952) contributed a systematic account of the ways in which clinical judgment is structured as it moves toward a specific prediction. Their analysis distinguishes five types of inference, which can be summarized as follows: (1) deductive--the derivation of a conclusion or assertion; (2) inductive--the derivation of a principle or continuum on the basis of common factors; (3) analogistic--the attribution of subsequent similarities to two phenomena which are similar in some initial respect; (4) eliminative--if there is a finite series of possibilities, A, B, and C, and the example cannot be classified as A or B, then it must be C; (5) postulational--one type of event is considered as if it were another kind of event.

The cues on which inferences are based are also classified by Sarbin and Taft into "classes" and "aspects."

Three classes of cues are: (1) analytic--readily communicable and easily identified; (2) pre-analytic--cues to which the inferring person responds but which are difficult to enumerate and locate; and (3) nonanalytic--the vague, poorly defined cognitive elements which arise from the self-perceptual field of the observer. "Aspects" include: (1) the locus of the cue (whether internal or external to the informing person); (2) the degree of accessibility to the inferrer's self-reactions; and (3) the manner (deliberate or automatic) in which the cues are used by the inferrer.[4]

How can the clinician proceed to build upon a predictive device? First of all, the type of inference used would be inductive, since the process of hypothesis formulation would be partially dependent upon the predictive information, i.e., he would be working outward from the information. Although the predictive device may not determine why a person succeeds or fails on parole, the information it contains can provide an initial point of inference. Secondly, the cue upon which the original inference is based is analytic, i.e., readily communicable and easily identified. Finally, the locus of the cue is external to the informing person (assuming the use of interviews) and the cues are used deliberately rather than automatically, by the inferrer.

---

[4]Much of this paragraph has been extracted almost verbatim from Gough (1962).

Table 2

The Process of Inference when Using
A Predictive Device in Conjunction with an Interview Sequence

Parole Supervisor's Questions and Ruminations	Offender's Responses
1. Rumination: Alcohol is indicative of parole failure and this man has a history of alcohol use.	
2. Question: "Maybe you can give me some idea as to why you use alcohol?"	3. Response: "Well, since my wife left about 5 years ago, I have been depressed at various times, and alcohol helps relieve my worries."
4. Question: "I notice on your recent arrest record that you were under the influence when you burglarized the residence."	5. Response: "Yeah, I sometimes feel pretty bold when I have a couple of drinks."
6. Rumination: It seems that he drinks to help forget his problems and then the alcohol acts as a stimulus to commit a crime.	
7. Question: "Do you have any present plans to get remarried?"	8. Response: "Well, yes, I am planning on marrying a woman I met last month."
9. Question: "When?"	10. Response: "Next month."
11. Rumination: Well, maybe his getting remarried will have an effect on his drinking habits and thus on his future criminal behavior.	

Table 2 provides an example, possibly oversimplified, of how an interviewer might derive further understanding from predictive information. By identifying mediating factors during the course of the interview, the case worker builds upon the information obtained to establish the stimulus-response chain: wife missing ——→ depression ——→ alcohol ——→ criminal activity. Although it is not resolved during the interview why alcohol leads to criminal activity (except for the general statement about reduced inhibitions), this step may not be essential as long as the probable cause of using alcohol is identified. The sequence of the scenario is important since (assuming that the chain is correct) the fact that the offender is to be married may have important implications for his success on parole. Although there is no way to determine the correctness of the hypothesis in this case, there is reason to believe that additional information derived during the course of the interview can lead to more accurate decision making.

Stouffer (1941) emphasized the importance of the clinician's ability to give more weight to a factor than it is given in an actuarial table. This observation is pertinent since a prediction device provides a static expectation of parole outcome which does not consider events that take place between the construction of the

device and the actual outcome. We have attempted to show that the clinician can build, on predictive information, hypotheses which bear upon the correctness of the original prediction. Consider an example derived from Table 3, which is the same as Table 1 except that the letter subgroup designations have been replaced with score values of a hypothetical parole prediction device. The higher the score, the greater the number of characteristics present which have been found to be associated with parole success.

Table 3
Hypothetical Parole Success Rates
for Five Groups of Offenders

Subgroup Score Intervals	Number		Percent	
	Success	Failure	Success	Failure
20-25	16	4	80%	20%
15-19	12	8	60%	40%
10-14	10	10	50%	50%
5- 9	8	12	40%	60%
1- 4	4	16	20%	80%

Let us assume that an offender, John Smith, obtains a score of 23, indicating that he has an 80 percent chance of succeeding on parole. However, during an interview John admits to prior opiate use, which is not reflected in his record. The case worker has uncovered an important piece of information which, although included in the formula for

calculating parole success probabilities, was mistakenly omitted in the calculation of John's chances on parole. Assuming (as much research has found) that a history of opiate use is related to parole failure, this additional information would decrease John's probability of parole success from 80 percent to, let us say, the fourth score category of 5-9 and an associated 40 percent probability of parole success. The interview has brought to light information which is of value to the actuarial prediction and should affect the prediction of John's probability of parole success.

Considering the great number and variety of individuals, it is easy to see how information important to the prediction may be overlooked by the actuarial device. Indeed, there are many kinds of events which can affect the probability of successful parole outcome but which occur only rarely. Considered singly, such factors may contribute heavily to prediction "misses." There are many ways in which the clinician might improve upon the actuarial device, provided that he exercises skill and care in obtaining additional information during the interview.

(D). Level III Rationale: Organizing Predictive Data - The DPPI Format

We have seen that prediction and explanation need not be incompatible if they are carefully merged and that the

process of hypothesis building can, in certain instances, follow from the use of predictive information.

An important assumption of the present study is that the case worker, confronted with a predictive instrument, will want to know what is not related, as well as what is related, to the criterion. Since certain kinds of information may be more valuable than others in predicting parole outcome, the case worker should be able to compare his subjective ideas of what is or is not related to the criterion with empirically derived findings. Thus, in addition to generating a specific predictive score, the DPPI seeks to provide a comparative framework within which the case worker can examine the accuracy of his own hunches.

Most "base expectancy" instruments provide little more than an indication of what is related to the dependent variable. If predictive information is presented in a simple summative style requiring only the tabulation of several different scores, we have a "cookbook" approach to prediction, which denies the clinician the opportunity for self-confrontation. A more useful instrument would be one which facilitated the comparison of assumed relationships between independent and dependent variables with demonstrated relationships. Rogers (1967) examined the degree of congruence between research results and the perceptions of correctional workers with respect to certain prediction

items. This kind of study indicates that parole prediction devices could provide, in addition to the predictive score, information for the comparison of correctional staff's hunches with empirically derived results.

Table 4 provides an example of a base expectancy table as it appears to the case worker. Although presented in a simple computational format, it does not address the issues of how prediction might facilitate understanding or how such a device might validate or modify subjective assumptions about predictive data. In contrast, the D.P.P.I. project assumes that data presentation and formatting can accomplish more than typically has been attempted. A prediction device should divide all relevant variables into associated groups (e.g., psychological, vocational, academic, etc.) and present predictive equations for each group. This framework facilitates case worker self-interrogation (e.g., does each group of variables contribute as much to the overall prediction as the case worker assumes?) and provides numerous leads for further examination and hypothesis building. Such an approach provides a "gestalt" or configuration of data that allows the comparison of hunches with empirically derived data and suggests possibilities for more intensive questioning. In addition, by taking any set of data and constructing case-history categories, the case worker can undertake analyses which appear to follow from the presentation.

Table 4
Calculation of Base Expectancy Raw Scores*

IF	ADD
No Prior Record	10
Limited Prior Record (not more than two jail or juvenile or one prison commitment). . .	4
Homicide, Assault, or Sex as most serious commitment offense under this serial number . .	6
NOT Burglary, Forgery, or NSF Checks as most serious commitment offense under this serial number	2
Age 30 or Older in year of release to parole	3
No History of Any Opiate Use	8
Original Commitment	1
Total Possible Score	34

Table 5[5] provides an example of how two groupings of data (vocational and offense information) would appear if the synthesis model were adopted. Two important characteristics should be noted. First, the predictive score is calculated for each individual for each dimension; and second, determining a score for each individual permits an

*Extracted from Gottfredson, D. M. A Shorthand Formula for Base Expectancies, CDC-BE-616, Research Division, California Department of Corrections, 1965 (Research Report No. 5).

[5]Presented with both figures is a bar indicating the base rate, or number of parolees who are actually successful on parole. For the present data, the base rate is 60.5%.

Table 5

Vocational Information

Add 1 if:

(1) Motivated for vocational training ____
(2) Employed six months or more ____

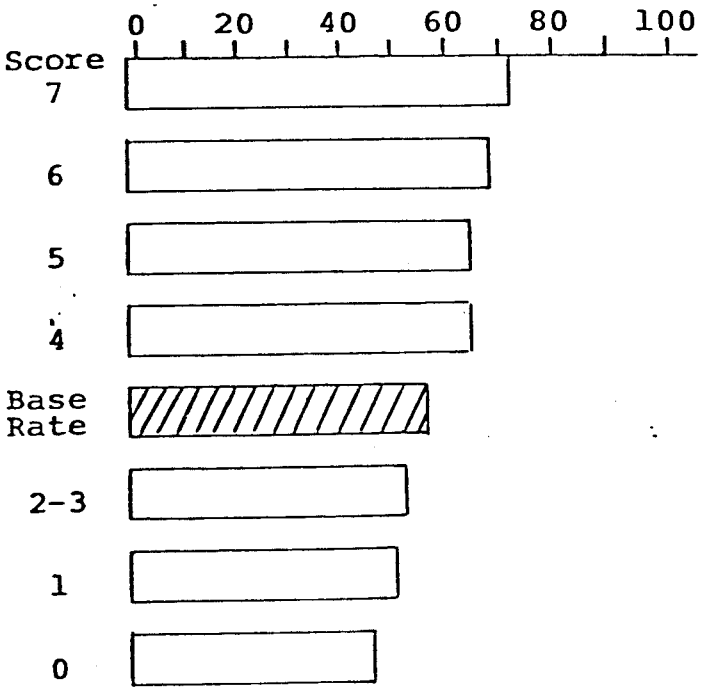
General Aptitude Test Battery

Add 1 if:

(1) General intelligence (G) score is 94 or above ____
(2) Verbal aptitude (V) score is 89 or above ____
(3) Clerical aptitude (Q) score is 97 or above ____
(4) Motor coordination (K) score is 101 or above ____
(5) Finger dexterity (F) score is 87 or above ____

TOTAL ____

Percent Success



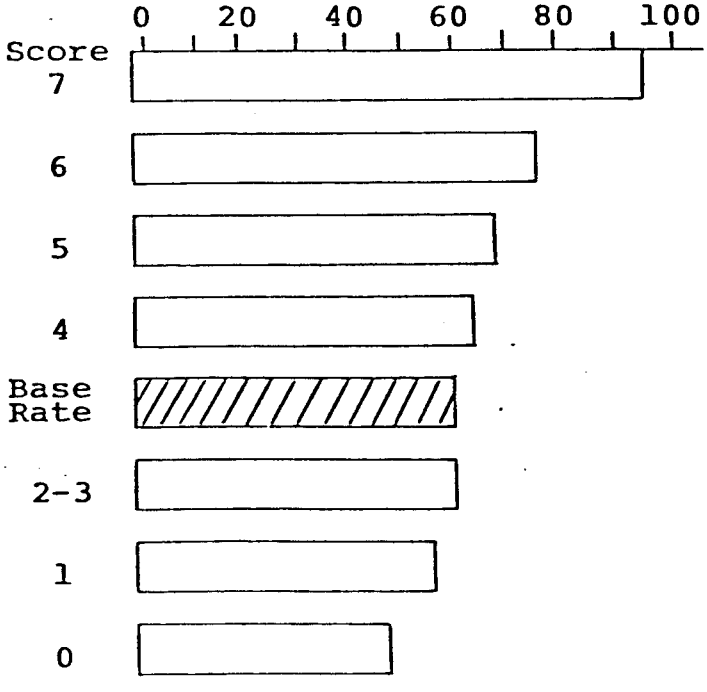
Offense Information

If present add 1:

(1) History of violence ____
(2) No history of escape ____
(3) Individual violence during offense ____
(4) Weapon used ____
(5) CYA violence history ____
(6) Diagnosed violence potential is least/mild ____
(7) Admission offense was against persons ____

TOTAL ____

Percent Success



estimation of the individual's probable parole outcome in relation to that dimension's ability to differentiate high and low risk groups. In the example in Table 5 it can be seen that offense information is better able to isolate extreme risk groups than is vocational information.

By placing an individual's score within one category and comparing the success rate of that category with those of other categories, one can see how much information is being derived in contrast to that derived from a simple base rate prediction. In looking at an individual's score, the case worker is asked to "weigh" the probabilistic determination of an individual's parole outcome in relation to each area of information and its separate ability to predict parole outcome. The instrument allows the calculation of an individual's score but also suggests the confrontation between what is assumed to be related to the criterion and what actually is related.[6]

IV. METHODOLOGY

A. Methodological Implications of the Different Techniques and Their Importance to the DPPI

Wallin (1941) distinguished three methods by which predictions may be made from the case study: (1) the case may be studied with reference to a series of factors known or assumed to be relevant to the prediction criterion; (2) the case may be classified typologically and a prediction

---

[6]Adapted from Rogers (1967).

made from the class; and (3) the case may be viewed as unique and an attempt made to identify idiosyncratic trends and project them into the future.

Wallin's observation is important to the present study because it provides a basis for distinguishing styles of perceiving patterns within the case-study setting. It might be appropriate to determine the degree of acceptance that a case worker with a particular decision-making style might show for an actuarial approach which stressed a similar style. The possibility that a certain prediction technique will be methodologically more compatible with some decision styles than with others adds a new component to the clinical synthesis model.

The first of Wallin's methods is not very different from what an actuary would do, given a certain type of predictive strategy. Two primary prediction techniques derive specific predictions from the presence or absence of a standard number of characteristics. The simplest actuarial method of this "constant factors" approach is the technique developed by Burgess (1928). This method proceeds by first identifying each variable that is related to the criterion and then assigning points on the basis of association between the predictor, or independent variable, and the criterion, or dependent variable. The presence or absence of characteristics related to the criterion determines the total number

of points assigned to an individual, who is then placed in a predicted score group with an associated rate of parole success.

A second constant factors prediction method is that of multiple regression, which, unlike the Burgess technique, weights variables according to their contribution to prediction of the criterion. Multiple linear regression produces an equation which expresses one (dependent) variable in terms of other (independent) variables, using the assumption that any relationships existing between the variables are linear.

From the case worker's perspective, the Burgess and multiple regression methods might provide somewhat different advantages. A case worker who feels that stability of the home situation, vocational training potential, and academic achievement, for example, are related to parole outcome may not consider them to be equally related to parole outcome, although the "weighting" of the variables is not made explicit. Thus the logic of weighting variables, if not the actual procedure, is the same in both the clinical and the statistical situations, and the case worker who uses the constant factors approach may find multiple regression, which helps him to formulate combined predictions involving several variables, to be more helpful than the Burgess method.

Wallin's second approach to prediction in an individual case follows from typological classification in which the individual prediction is made from the associated class. This procedure is similar to the process by which predictive attribute analysis (PAA) derives its prediction. This technique results in the identification of relatively homogeneous subgroups, each with an associated configuration of characteristics. PAA, a divisive, hierarchical method of clustering individuals, proceeds by repeatedly dividing groups in two, making a "tree." The resulting configuration forms terminal groups consisting of different combinations of characteristics with different probabilities of parole success. Instead of deriving each individual's score from a number of predetermined factors, PAA proceeds by determining the characteristics an individual shares with available subgroups. A case worker using predictive attribute analysis must place an individual into a subgroup within an empirically derived typology before his associated parole outcome can be estimated. Again, it is possible to view this technique as more compatible with certain case-study procedures than with others.

Wallin's third approach to case-study prediction involves the identification of idiosyncratic characteristics (e.g., observed moods or emotional states), with which actuarial prediction has little concern. Before assuming, however,

that there is no actuarial prediction technique similar to this method, it may be instructive to recall Allport's (1940) distinction between "actuarial" and "individual" predictions, i.e., that the prediction for a class and the prediction for an individual are different procedures. While Wallin may have identified a case work process which, from the actuary's perspective, excludes the statistical table, he does not state that predictions for a class and for an individual are different concerns, but only that certain kinds of information cannot be used in actuarial prediction.

To further investigate whether one of the above-mentioned actuarial methods might provide the synthesis framework, we must examine certain characteristics of each method, following from mathematical considerations, which might affect the value of the clinical synthesis model for different decision styles. There are a number of methodological considerations which bear upon prediction and thus upon the clinical interpretation of predictive information.

One important consideration in the selection of a prediction technique is whether it takes into account predictor intercorrelation, a term which refers to the fact that the predictors are themselves interrelated. Assume that a case worker wishes to estimate the likelihood of successful parole for a ward. Assume the ward has a history of violence and that wards with such histories have a parole

success rate 6% higher than the overall group (base rate). Assume also that the ward is 20 years old and that boys of this age have a parole success rate 4% higher than the base rate. The case worker might assume that the ward's chances for successful parole outcome are 10% (6% + 4%) above the base rate. Since previous research has found that age and violence history are related, this assumption would be incorrect. Because of their intercorrelation, the contributions of these two variables are not summative; although each variable contributes some information to the prediction, their combined contribution is somewhat less than a simple total. The case worker may be misled if he either falsely assumes the prediction to be summative or relies on an actuarial method which does not account for predictor correlation. Not only might an inaccurate prediction be made but also other leads identified during the course of an interview might be falsely interpreted because of the initial error. Predictor intercorrelation thus is an important consideration in the synthesis model.

A second important methodological consideration is sample overfitting, which results from the assumption that characteristics shown to be related to parole outcome in one population are similarly related in a second population. Since two groups of subjects are never identical, we can expect some variation in results between groups. For example,

if we find in one group that older wards do better on parole we might reasonably expect older wards to do better in a second group. However, the strength of the relationship is not necessarily the same for all groups. Those prediction methods which fit the sample most closely run the greatest risk of overestimating the extent to which new data can be explained by relationships found in the old.

Another issue to be considered is whether the predictive method divides or partitions the sample into subgroups of offenders. Some prediction techniques partition the sample while others proceed by scoring each individual on a pre-set number of variables. Recalling the comparison of the constant factors method of prediction and the typological method, we can see that this characteristic is crucial to Wallin's (1941) distinction concerning decision styles, and the clinical synthesis method is sensitive to these styles.

A fourth consideration in the comparison of predictive methods is the assumption of the linearity of a relationship. If two variables are perfectly related, their relationship appears as a straight line. From the case-study perspective, the issue of linearity may be important because curvilinear relationships often are not tested with most prediction techniques. For example, a clinician may assume that parole success rates are higher for a medium intelligence group than for either dull-normal or superior groups. Unfortunately, the correctness of this assumed relationship cannot easily

be tested by multiple regression or any other commonly used method because the relationship assumed is not linear. Since they do not allow feedback for revision of such judgments, actuarial devices not only fail to provide important information but may also, in some instances, be misleading. Important relationships may be missed because of the inability of prediction methods to identify non-linear associations.

Finally, some methods which have been used in prediction studies are not, in the strictest sense, prediction methods. These methods do not employ a dependent variable; rather, they attempt to form subgroups of the total population whose members are similar simultaneously on several variables. The use of these methods for prediction involves the reasonable assumption that people who are members of groups which are similar in general will be similar on a specific variable--outcome.

It is possible to rate prediction methods on their ability to deal with predictor intercorrelation, sample overfitting, sample subdivision, and linearity. Table 6 provides a comparison of four methods--multiple regression, Burgess, association analysis, and predictive attribute analysis--with respect to these issues. As Table 6 shows, no one method is most desirable on all dimensions. There must therefore be a "trade-off" among desirable and undesirable features in all instances. As an alternative, we could

utilize Wallin's (1941) classification and conclude that the best method in a given situation is the one most compatible with the particular decision-making style. While methodological considerations are important, they should not obscure the more basic goal of this study, i.e., to develop a synthesis of actuarial and clinical concerns. Since the proposed model is to be integrated into the case-study context, selection of the "best" instrument should remain in the hands of the practitioner.

Table 6
Ratings of Four Prediction Methods
on Five Characteristics

Method	Accounts for Predictor Inter-correlation?	Sub-divides sample?	Tendency for over-fitting	Assumes Linear Relation-ships	Uses Dependent Variable?
Burgess	No	No	Low	*	Yes
Multiple Regression	Yes	No	Moderate-High	Yes	Yes
Association Analysis	Yes	Yes	Low-Moderate	*	No
Predictive Attribute Analysis	Yes	Yes	High	*	Yes

*Since all variables are dichotomized, non-linear relationships are not possible.

B. Comparative Prediction Analysis

The DPPI project proposed to compare the predictive efficiency and appropriateness to the DPPI clinical format

of several prediction techniques. The "trade-off" between predictive power and meaningful input to the case-study processes was to determine the optimum predictive strategy.

The comparative prediction analysis was composed of the following steps:

(1) Several subsets were formed from the over 150 variables available. These subsets were selected to represent various dimensions presumed to be relevant to case-study processes. These dimensions were: individual case history, psychological, social psychological, vocational, intelligence, academic, and offense-specific.

(2) Two prediction techniques were applied using the variables within each dimension. This resulted in seven prediction equations, each based on variables from only one dimension. The validity of the predictors was tested by applying each equation to a validation sample. The purpose here was to explore the predictive power of each of the seven subsets of variables.

(3) Four prediction models were applied to the construction and validation samples but predictor variables were drawn from all dimensions rather than exclusively from one dimension. The result was one composite prediction equation for each method. The purpose here was to explore the predictive power of the four methods and to provide a bench mark against which the predictions developed from the subsets

could be compared.

(4) An analysis was performed. First, the results obtained when predictions were developed within dimensions were compared. Second, various prediction methods were compared on their ability to predict outcome when all variables were used. Finally, the results of step 2 were compared to those of step 3 to assess the loss of predictive power which resulted from limiting predictors to clinically relevant dimensions.

Predictive Methods Used in This Study

For comparison, prediction strategies were selected which differed on the characteristics of predictor intercorrelation, linearity, additivity, sample subdivision, focus on a dependent variable, and sample overfitting.[7] The techniques selected and summaries of their characteristics follow:

(1) Burgess Technique - This method was developed by Burgess (1928) and was selected because of its simplicity and promise of good validity. Each factor which is related to the criterion is identified. Variables which are significantly related to the criterion are used to develop a predictive score for each individual in the sample. Each person's predictive score is the number of attributes he possesses which are positively related to the criterion.

---

[7]See pages 55-59.

(2) Multiple Regression - Multiple regression produces an equation which expresses one variable (the "dependent" variable) in terms of others (the "independent" variables). Since there is more than one independent variable in the regression equation, the contribution of each variable to prediction of the dependent variable is expressed as a "weight" which varies with the extent of the relationship between the independent and dependent variables. In the regression procedures used in this study, independent variables were entered into the equation in a step-wise fashion. Predictive scores were developed by computing the predicted value for each individual.

(3) Association Analysis - Association analysis is a non-linear configural method of classifying individuals into groups. Originally developed by Williams and Lambert (1959), it is widely used in the physical and social sciences. For association analysis, the data on individuals must all be in attribute form (i.e., all variables must be dichotomized) and the individuals are grouped according to whether they possess or lack certain attributes. The technique proceeds by successively dividing the original sample into two parts until a stopping rule is satisfied. All individuals are thus classified into sets, which are at the terminal points of the "tree." Each successive division is made on the attribute which is most related to all other attributes

which describe the members of the group. The method is primarily descriptive rather than predictive but the resulting classifications produce relatively homogeneous groups which may have different parole success rates.

(4) Predictive Attribute Analysis - Developed by MacNaughton-Smith (1965), predictive attribute analysis resembles association analysis in that it is a configural method of clustering individuals. The clustering is directed throughout toward the prediction of a specified criterion, however. All variables, including the criterion, are converted to attribute form and the sample is repeatedly divided into two, making a "tree." The attribute chosen for splitting at any stage is the one that has the highest relationship with the criterion.

Techniques Used for the Comparison of Results

A number of techniques were used to compare the results obtained with different prediction methods. These techniques measure either variation explained or the classificatory accuracy of a prediction table.

The variation based methods, R^2, r^2, and eta^2, represent measures of the proportion of the total variation of the dependent variable which is explained by the prediction. The different measures are appropriate to different situations. R^2, the Coefficient of Determination, is used to summarize the explanatory power of all the predictor variables

in a multiple regression equation. The explanatory power of a prediction developed using the Burgess method can be summarized by r^2; each element in the sample is scored according to the Burgess device and these scores are correlated with the actual outcome, the resulting correlation coefficient, when squared, represents the proportion of variation in the dependent variable which is explained by the Burgess device. When applying a prediction equation developed using multiple regression to a validation sample, a similar procedure is used: predicted scores are calculated for each sample element; these are correlated with outcome and the resulting coefficient is squared to yield r^2, which indicates the proportion of variation of the outcome "explained" by the predicted scores.

Eta^2 also represents the proportion of variation explained by prediction but is appropriate for methods (such as association analysis and predictive attribute analysis) which classify sample elements into groups rather than yielding predicted values for each element in the sample.

For simplicity of exposition, the term "r^2" is used to represent each of these variation based measures in the following discussion.

Mean Cost Rating (MCR), developed by Duncan et al (1953, 1955), is designed to reflect the extent to which a prediction device has successfully classified individuals into categories of diverse success rates. The MCR does not assume normality,

continuity, or equally spaced scores. MCR coefficients were calculated for all four predictive methods--the two configural methods as well as Burgess and multiple regression--thus allowing a straightforward comparison of classification error.

C. Field Evaluation

Since an important objective of the DPPI project was the development of a model of a field-relevant instrument - i.e., one that would be useful to the practitioner in his decision making within the case-study process - the design of the project envisioned the involvement of field staff at various levels in assessing at least the "face applicability" of the model.

Within the constraints of time and budget, maximum statewide field representativeness was to be sought. Various types of staff with different levels of experience and education were to be interviewed. Since until more current cross-validation studies had been carried out it could not be assumed that the DPPI project would provide a usable tool, the primary measure of field relevance at this early stage had to be practitioners' attitudes toward the instrument. Because no attitude device assessing field opinion relevant to this project could be found, a questionnaire was constructed to obtain opinions on three major areas of concern to the evaluation: (1) Is the prediction of future behavior useful in case-study work? (2) Could the DPPI, in any of

its alternate forms, be useful in the case-study process? How? (3) Are the proposed formats of the instrument understandable?

The field evaluation component was designed as a two-phase task overlapping the development of the alternative DPPI formats. Evaluation was to be both ongoing or formative and summative, with initial introduction to the field followed by modification of the model along suggested lines and re-introduction of a modified instrument at a later date.

The validity of two basic assumptions underlying the development of the DPPI model was to be tested by the field evaluation. As we have seen, these assumptions, which guided the efforts to create a clinically useful prediction tool, were: (1) that the organization of predictive data into case-study dimensions can facilitate self-interrogation by the case-worker; and (2) that the use of different prediction methods will produce different formats, which coincide with different decision-making styles. The different prediction techniques and their associated formats were to be assessed in terms of simplicity, comprehensibility, and utility (appropriateness to styles of decision making).

The original design of this phase of the project had to be substantially revised because of time limitations of staff at the California Youth Authority field sites.

Telephone Survey

Although the full two-phase evaluation could not be undertaken, preliminary contacts were made with field staff to obtain an idea of their familiarity with parole prediction research, their attitudes toward prediction instruments, and the sources of information on which they relied in their daily decision making. It was believed that the information gained from these preliminary contacts also would be useful in designing the full-field questionnaire and interview schedule.

An evaluation staff member set up telephone interviews of about twenty minutes each with a small (N = 10) sample of California Youth Authority personnel involved with release planning and parole supervision. A brief field questionnaire covered such questions as: (1) What type of outcome criterion is most important to your decision making? (2) What sources of information do you use when making decisions? (3) Do you rely primarily on objective information or do you use both objective and subjective (interview-type) sources? (4) Do you think statistical prediction of parole outcome is useful? (5) Do you presently use or have you used any parole prediction instrument? (6) Do you feel that typologies are useful in your case-work contacts?

Interviews and Questionnaires

A full-field questionnaire was designed to be administered

by evaluation staff in the course of an interview so that staff could explain the project in detail, respond to any questions about its design and purposes, and note any freely given comments of field personnel that might indicate their general attitude toward statistical prediction of their specific reactions to the development of an instrument such as that envisioned by the DPPI project. Questionnaires on the three formats designed for the DPPI (Burgess, multiple regression, and predictive attribute) also were administered where possible.

Efforts were made to obtain interviews with personnel at four points in the correctional sequence: pre-institutional staff (at a reception center-clinic), correctional institution staff, parole agents, and parole board members. Parole board members were interviewed briefly as a group, while fairly intensive field contacts (from one to two hours in length) were concluded with personnel at the other three levels. The personnel interviewed represented urban, suburban, and rural areas of Northern California.

V. RESULTS AND CONCLUSIONS

A. Comparative Prediction Analysis

i. Results

Predictions from Dimensions

As mentioned above, the 150 variables in the data base were divided into seven "dimensions," which represented

groups of variables presumed to have a common relevance to case-study concerns. These dimensions were:

Individual case history

Psychological factors

Social psychological factors

Vocational factors

Intelligence factors

Academic factors

Offense-specific information.

The variables in each of these dimensions were used to develop predictions using two methods--multiple regression and the Burgess technique. These methods were selected for this analysis because they differ substantially on a number of properties and because exploratory analysis had indicated that the configural methods behaved poorly when applied to the dimensions with a large number of highly interrelated variables.

The purpose of this analysis was (1) to compare the dimensions on their relative ability to predict outcome and (2) to develop the information needed to estimate the loss of predictive power which results from restricting predictor variables to clinically relevant dimensions.

The results of the predictions developed from the seven dimensions were compared. The MCR and r^2 coefficients for each of the seven dimensions for the multiple regression and

the Burgess techniques are reported in Table 7 below.

Table 7
MCR and r^2 Coefficients for Multiple Regression and Burgess Predictions from Seven Dimensions

Dimension	Mean Cost Rating (MCR)		r^2	
	Construction	Validation	Construction	Validation
Individual Case History				
Multiple Regression	.212	.129	.0424	.0058
Burgess	.215	.139	.0357	.0130
Psychological Factors[a]				
Multiple Regression	.097	.005	.0072	.0000
Burgess	.058	-.003	.0023	.0001
Social Psychological Factors[b]				
Multiple Regression	.130	.022	.0159	.0001
Burgess	.093	.007	.0052	.0001
Vocational Factors[c]				
Multiple Regression	.088	.014	.0085	.0004
Burgess	.090	.016	.0062	.0001
Intelligence Factors[d]				
Multiple Regression	.043	.034	.0015	.0008
Burgess	*	*	*	*
Academic Factors[e]				
Multiple Regression	.090	.050	.0077	.0028
Burgess	.080	.024	.0060	.0005
Offense-Specific Information[f]				
Multiple Regression	.132	.099	.0182	.0062
Burgess	.104	.077	.0102	.0058

*No Burgess prediction was developed from Intelligence factors since no variable in this dimension was significantly related to outcome.

[a]consist of subscale scores from the Minnesota Multiphasic Personality Inventory

[b]consist of subscale scores from the California Psychological Inventory

[c]consist of such items as length of employment, motivation for vocational training, and several subscale scores from the General Aptitude Test Battery

[d]consist of various intelligence measures

[e]consist of such items as age left school, highest grade claimed, and California Achievement Test scores

[f]consists of seven variables which relate to the ward's commitment offense.

Two observations emerge clearly from a review of the data presented in Table 7. First, the predictive power obtained from even the best dimension is disappointingly low. This observation will be discussed further in the next section (Overall Prediction and Comparative Analysis). Second, there is substantial variation in the power of predictions developed from the seven dimensions.

Table 8
Ranking of Dimensions on the Basis of MCR and r^2 Coefficients

Ranking	Highest r^2 Obtained (Validation)		Highest MCR Obtained (Validation)	
	Dimension	Coefficient	Dimension	Coefficient
1	Offense Specific	.0062	Individual Case History	.138
2	Individual Case History	.0058	Offense Specific	.099
3	Academic	.0028	Academic	.050
4	Intelligence	.0008	Intelligence	.034
5	Vocational	.0004	Social Psych.	.022
6	Social Psych.	.0001	Vocational	.014
7	Psycho-logical	.0000	Psycho-logical	.005

Table 8 provides a list of dimensions ordered in terms of both MCR and r^2 coefficients. Individual case history and offense-specific information achieved the best r^2 values, although neither dimension explains more than 1% of the variance on validation. These findings are not surprising:

most previously developed predictive devices have found background history and offense information to be the best predictors of parole outcome. None of the remaining dimensions explained more than .28% of outcome variation. Interestingly, the majority of the variables which make up the remaining subsets are based either on psychometric, intelligence, and achievement test scores or on such measures as the subjective ratings of case workers. This finding suggests that static background characteristics provide the most stable and reliable indicators of parole adjustment. The MCR coefficients provide a similar yet not identical ranking of dimensions. Both the individual case history and offense-specific dimensions have higher coefficients than the remaining subsets, but the MCR's are not impressive.

A comparison between the multiple regression and Burgess methods of prediction using the seven dimensions is provided in Table 9. Four columns report the methods which produced the highest coefficients (r^2 and MCR) for construction and validation samples. The "least difference" columns name the method achieving the smallest change between the construction and validation sample coefficients. This is a simple indicator of shrinkage. The data in Table 9 indicate that although multiple regression generally evidences better "data fitting" properties, it is more subject to overfitting and consequent shrinkage than is the Burgess method. This fact is also

Table 9
A Comparison of Multiple Regression and Burgess Methods of Prediction Across Seven Variable Subsets

Dimension	r^2			MCR		
	Highest r_1^2 (Construction)	Highest r_2^2 (Validation)	Least Difference (r_1^2-r_2^2)	Highest MCR_1 (Construction)	Highest MCR_2 (Validation)	Least Difference (MCR_1^2-MCR_2)
Individual case history	Regression	Burgess	Burgess	Regression	Burgess	Burgess
Offense specific	Regression	Burgess	Burgess	Regression	Regression	Burgess
Academic	Regression	Regression	Burgess	Regression	Regression	Regression
Intelligence	*	*	*	*	*	*
Vocational	Regression	Regression	Burgess	Burgess	Burgess	**
Social Psychological	Regression	**	Burgess	Regression	Regression	Burgess
Psychological	Regression	Burgess	Burgess	Regression	Regression	Burgess

*The Burgess method was not applied to the dimension.

**The values are identical in these cases.

indicated in the comparison of the validation coefficients for both r^2 and MCR, which finds the Burgess technique superior in five out of eleven cases. This finding indicates that, in comparison with the Burgess method, multiple regression tends to "overfit" the construction sample so that applications of its predictions to the validation sample are relatively less effective.

Overall Prediction and Comparative Analysis

Another major undertaking of the project was the comparison of several techniques in their ability to predict the criterion without regard for the clinically relevant dimensions. This analysis was performed with the four prediction methods previously identified: multiple regression, Burgess, association analysis, and predictive attribute analysis. The predictive efficiency of the methods was assessed by comparing Eta^2 (as a measure of variation explained by configural prediction methods) or r^2 and MCR. Table 10 reports the values used in this comparison.

Table 10
Comparison of Prediction Methods
Using MCR, r^2 and Eta^2 Coefficients

Prediction Method	MCR		r^2 or Eta^2	
	Construction	Validation	Construction	Validation
Multiple Regression	.289	.139	.0655	.0083
Burgess	.169	.058	.0219	.0024
Predictive Attribute	.324	.167	.0697	.0220
Association Analysis	.158	.129	.0190	.0129

In this table, construction and validation Mean Cost Ratings and r^2 or Eta^2 are presented for each of the prediction techniques. Predictive attribute analysis displayed the highest predictive power of any of the methods on both construction and validation and explained 2.2% of the variation in parole performance for the validation sample. Association analysis explained less of the variation in outcome on construction than any other method (less than 2%)--a result which was expected since association analysis is not fundamentally a predictive technique. Yet, also as expected, its predictive power shrank the least between construction and validation and it emerged as the second most powerful technique on validation.

The regression technique performed nearly as well on construction as predictive attribute analysis, explaining over 6.5% of the variation; however, on validation this method accounted for only .8% of the variation. The Burgess method explained 2.1% of the criterion variation in the construction sample, although this fell to .2% on validation. Of the four methods, the configural approaches seem to perform somewhat better than the others.

In Table 11 each method is ranked with regard to the highest r^2, Eta^2, or MCR coefficients obtained in construction and validation. Also, to provide a simple measure of shrinkage, the differences between construction and validation

Table 11
Comparison of Ability of Four Prediction Methods
to Predict Parole Outcome

(N = 4,146)

Rank	r^2			MCR		
	Highest r_1^2 or Eta_1^2 (Construction)	Highest r_2^2 or Eta_2^2 (Validation)	Least Difference $r_1^2-r_2^2$ or $Eta_1^2-Eta_2^2$	Highest MCR_1 (Construction)	Highest MCR_2 (Validation)	Least Difference MCR_1-MCR_2
1	Predictive Attribute Analysis	Predictive Attribute Analysis	Association Analysis	Predictive Attribute Analysis	Predictive Attribute Analysis	Association Analysis
2	Multiple Regression	Association Analysis	Burgess	Multiple Regression	Multiple Regression	Burgess
3	Burgess	Multiple Regression	Predictive Attribute Analysis	Burgess	Association Analysis	Multiple Regression
4	Association Analysis	Burgess	Multiple Regression	Association Analysis	Burgess	Predictive Attribute Analysis

coefficients are ranked. As noted in Table 11, the "least difference" columns show that association analysis achieves the greatest stability from construction to validation. Since association analysis is not basically a prediction method and is therefore least susceptible to overfitting, this result could be expected. This also explains its poor performance (the lowest MCR and r^2) on construction.

Multiple regression performed nearly as well as the predictive attribute method, achieving the second highest r^2 and MCR on construction. However, due to this method's tendency to overfitting, there is a sizable reduction in the coefficients achieved on validation. Overfitting is further demonstrated by this method's relatively poor showing in the least difference columns, where it ranked fourth on the r^2 measure and third on the MCR measure column.

The Burgess technique performed least efficiently of all techniques. Achieving relatively poor rankings on both r^2 and MCR measures, the Burgess method is consistently lower than both predictive attribute analysis and multiple regression. In terms of the shrinkage measures, the Burgess method ranked second only to association analysis. This result was expected since the Burgess technique is not prone to substantial overfitting.

Predictive attribute analysis performed better than any other method, achieving the highest r^2 and MCR coefficients

in both construction and validation samples. As expected, this method performed somewhat less efficiently on the shrinkage measures.

It is interesting to note that the configural techniques have dominated the comparative analysis. In terms of both highest r^2 and MCR values and least shrinkage, predictive attribute and association analysis perform better than either the multiple regression or the Burgess technique.

However interesting the comparison of methods may be, the most noteworthy feature of the analysis is the generally poor level of predictive power achieved. Compared with the results of similar studies, the results presented here evidence relatively little predictive power: the best prediction (with predictive attribute analysis) explained only 6.9% of the criterion variance on construction and 2.2% on validation. Also noted was the poor performance of the regression analysis, which is surprising when its superior data fitting capabilities are considered. The fact that both configural techniques produced better predictions in the validation sample than either of the additive techniques further suggests that substantial nonadditive relationships exist in the data.

Nonadditive effects could be caused by "masked heterogeneity." This term refers to the result of collapsing heterogeneous subgroups into a single population, thus masking

the possibility that different predictors are relevant for different subgroups. Although analysis of this possibility was not targeted for completion during the project period, an attempt was made to determine whether masked heterogeneity might be responsible for the poor predictive power obtained.

The entire study population was subdivided on the basis of commitment offense into three groups: person offenses, property offenses, and drug or alcohol offenses. The construction sample for property offenses consisted of 1,950 cases, while the validation sample consisted of 476 cases. Construction and validation samples for property offenses consisted of 709 and 148; for alcohol/drug offenses these were 329 and 78, respectively. To test the assumption that different predictors might be relevant to different subgroups, a multiple regression equation was developed for each subgroup.

The results for the three subgroups are reported in Table 12. The proportion of variation explained on construction for person, property, and drug/alcohol offenses are 6.8%, 5%, and 11.4% respectively. MCR's for construction subgroups are all fairly high, although on validation these are substantially reduced. On validation, the r^2 coefficients are reduced dramatically in the case of person and drug/alcohol offenses, with the reduction not as great for property offenses. These reductions should be interpreted with caution, however, since the reduced r^2 validation coefficients may be

due to the small size of the validation samples (148 and 78) for these two groups. Unfortunately, the small size of the validation samples disallows a determination of the amount of the reduction that is due to sampling error and the amount that is due to shrinkage. Nevertheless, the substantial r^2 and MCR coefficients for person and property offenses on construction indicate that the subdivision of sample populations into subtypes based on commitment offense may be a promising area for further research.

Table 12
Prediction Results Using Multiple Regression to Predict Parole Outcome for Three Subgroups

Subgroup	MCR		r^2	
	Construction	Validation	Construction	Validation
Person	.298	.032	.0681	.0003
Property	.226	.149	.0502	.0190
Drugs/ Alcohol	.394	.068	.1142	.0006

A similar procedure was followed with respect to admission status and race. The sample was divided into subgroups based on these variables and individual prediction equations developed for each subgroup. The results of this analysis are summarized in Table 13 below.

Table 13
Summary of Subgroup Predictions

Subgroup	MCR		r^2	
	Construction	Validation	Construction	Validation
Overall	.289	.139	.0655	.0083
Admission Status				
First Admission	.212	.137	.0421	.0168
Second and Subsequent Admission	.228	.060	.0567	.0045
Race				
Black wards	.281	.161	.0618	.0177
Mexican-American wards	.287	-.002	.0675	.0011
White wards	.391	.254	.0950	.0432

As is indicated by the coefficients in Table 13 the results of this analysis were similar to those obtained in the offense subgroup analysis. The power of the predictions in the construction sample lends some support to the hypothesis of "masked heterogeneity," but the substantial shrinkage on validation raises the question of whether the results reflect important relationships in the data or merely sampling error. Resolution of this issue must await further research.

Comparison of Predictions from Dimensions with Overall Predictions

One objective of this study was to determine whether restricting predictors to clinically relevant dimensions would substantially reduce predictive power. Above, the

development of predictions from seven such dimensions was described, as were the results of the effort to develop the most powerful prediction device without regard to the clinical relevance of the predictors.

The best overall prediction was obtained using the predictive attribute technique, which produced r^2's of 7% on construction and 2.2% on validation and MCR's of .324 and .167 on construction and validation, respectively. These coefficients indicate greater predictive power than was achieved using variables from any single clinically relevant dimension. The most powerful prediction developed in the dimensional analysis produced r^2's of 3.6% on construction and 1.3% on validation with MCR's of .215 and .139. These findings suggest that substantial predictive power is lost when potential predictors are limited to any single clinically relevant dimension. This conclusion must be tempered, however, by the fact that the predictive power found in this study is considerably lower than has been found in similar prediction studies, suggesting that attempts to develop predictions from clinically relevant dimensions should not be abandoned unless this finding is replicated in other studies.

ii. Conclusions

The following findings emerge from the results reported

above:

(1) Objective information, particularly that related to offense and background, predicted better than subjective judgments, including such variables as workshop and vocational recommendations. Offense and background information predicted parole outcome better than subsets consisting of psychological, aptitude, and intelligence test scores. Much prior research has noted the predictive power of criminal record (e.g., Ohlin, 1951; Glaser, 1954; Mannheim and Wilkins, 1955; Babst, 1964; and Simon, 1971). These data are usually hard, objective, and (depending on the accuracy of the records) reliable measures.

(2) Predictions developed by selecting variables without regard for the clinically relevant dimensions were more powerful than those developed by using variables exclusively from any single dimension. This finding implies that attempts to integrate actuarial prediction into the case-study process by restricting potential predictors is likely to result in a less than optimum actuarial device. This suboptimality must, of course, be balanced against the benefits achieved by inserting predictive information into the case-study setting.

(3) Predictive attribute analysis and association analysis, configural prediction approaches which do not assume additive relationships, performed better than did multiple regression and the Burgess technique, which do assume additive

relationships. This finding supports the suggestion of Grygier (1966) and others that many of the relationships in data such as ours are non-additive.

(4) The hypothesis of "masked heterogeneity," which could explain the presence of non-additive relationships in our data, was supported by the r^2's obtained when predictions were developed on subsets of the original data. We remain uncertain whether to attribute the shrinkage of r^2's on validation to sampling error or to unstable underlying relationships. The findings suggest that a procedure described by Babst, Gottfredson, and Ballard (1968) might be usefully applied to prediction problems. This procedure would involve the application of a configural method to reduce non-additive relationships, followed by multiple regression within the subgroups.

(5) The predictive power of the instruments developed in our analysis was poor. While other prediction studies report validation MCR's of from .25 to .44, .17 was the best achieved with our data. A plausible explanation for this finding is that our techniques were inadequate to reveal the predictive relationships in the data. While it is not possible to conclusively reject this possibility, the major approaches to prediction were represented in the techniques used. It therefore seems likely that if hidden relationships exist in these data they cannot be discovered using common prediction

methods. Another possibility is that for the group of wards represented in our sample, parole outcome is not predictable from the variables in our data. This might be the case because for this group post-intake events may strongly influence parole outcome. If this were the case, our data, collected at intake, would have little predictive power. It is also possible that important variables, known at intake, were omitted from the data.

Since the research described here employed an ex post facto design, the potential predictors were limited to the intake variables which were in the data set. For this reason it was not possible to determine whether additional intake variables or variables regarding post-intake events would appreciably increase predictive power.

B. Field Evaluation

i. Results

Telephone Survey

Responses to the question: "Is parole success or failure a crucial consideration in your decision making?" included "yes," "sometimes," "to some extent," "no," and "not at all." Only three respondents indicated that success on parole was a primary consideration. Other criteria considered important included the ward's self-esteem and satisfaction with life, adjustment in the community, improved social functioning,

and "stability."

Seven of the respondents indicated that in making decisions they relied on both objective and subjective information, while three responded that they used only objective information. However, the latter three included in their list of primary information sources: home situation, offender attitudes, case history, written reports by professional staff, and recommendations by agents or institutions, and all reported that interviews with the offender were an integral part of their decision-making process, indicating that they relied on subjective as well as objective information sources.

Interviews and Questionnaires

The findings of the field evaluation interview phase were analyzed separately for institutional staff, parole field staff, reception center staff, and parole board members in order to distinguish differences in orientation (possibly resulting from different degrees of concern with the criterion of parole outcome) and differences in reaction to parole prediction in general and to the DPPI in particular.

Institutional Staff

Of the eleven institutional staff members interviewed, only one agreed to complete the separate format comparison

questionnaire. This respondent strongly favored the Burgess format, feeling that the multiple regression and predictive attribute formats were too complicated for staff use in most facilities. He felt that the Burgess format allowed easy calculation of a ward's probable parole success or failure, fairly easy identification of subgroups with differing parole success rates, and fairly easy comparison of each information dimension's ability to predict parole outcome. He did not feel that the Burgess format attempted to fit too much information into too little space and agreed that this device could be helpful in decision making. He agreed also that the use of information dimensions facilitated the comparison of actual prediction results with subjective feelings about the relationship between a variable and parole outcome.

Of the eleven institutional staff interviews, only nine produced results for the professional questionnaire. Additional information was obtained through intensive questioning of respondents by the interviewer. Only two respondents to the professional questionnaire were aware of statistical prediction research in the field of corrections. Six respondents believed that prediction research might some day be helpful, while one thought it would not and two were uncertain of its potential utility.

Of the eleven staff members interviewed, six reacted in a generally positive manner to the concept of the DPPI and

indicated that, with validation and a few modifications, the clinical synthesis model would be useful in their case work. One respondent's attitude was negative and four expressed mixed feelings. Most of the eleven staff members expressed some skepticism about parole prediction in general and the accuracy of any prediction instrument. Many of them emphasized that the validity of any instrument would have to be clearly demonstrated before they would consider using it. Concerning format, institutional staff were almost unanimously in favor of the utmost simplicity if the instrument were to be used.

The orientation of the majority of institutional staff could be described as intuitive. Only one person indicated that he relied primarily on objective factors in decision making, five described themselves as primarily intuitive, and the rest stated that they relied on both objective and subjective or intuitive information.

When asked about potential uses for the instrument, four respondents suggested that they would use it to attempt to substantiate their intuitive judgments. Other possible uses suggested included evaluating a ward, classifying, and programming. Four of the six respondents who stressed programming emphasized that the ward should be involved in developing his own program with the help of the DPPI. Six respondents felt that the ward should have access to DPPI

results; one felt that he should not, because of the "self-fulfilling prophecy." However, only two respondents felt that the self-fulfilling prophecy would be a serious problem in the use of the DPPI.

Some of the additional comments on the DPPI were: that more social/environmental factors should be considered (five respondents); that progress in the institution does not indicate success on parole and that parole outcome does not reflect the success of institutional treatment (seven respondents); that more attention should be directed to change factors in the institution and that an index should suggest potentials, not just static scores, with a longitudinal view of progress in each dimension (four respondents); that such an instrument needs constant updating and improvement (one respondent); and that a list of probable implications of the data should be added (one respondent).

The professional questionnaire included an item on the kinds of data used by field staff in their decision making. Responses (which may overlap somewhat) are presented in Table 14.

Asked what kinds of information not presently available they would like to see included in a ward's file, respondents suggested: a Base Expectancy score or other probability evaluation; work background and aptitude test scores; MMPI and CPI scores; interview with family and ward together at the

Table 14
Information Used by Institutional Staff

Type of Information	Frequency of Mention
Interview with Ward	5
Clinic Summary	4
Test Scores	4
Psychiatric Reports	3
Staff Observations	3
Home Visit, Parent Interviews	3
Offense History & Pattern	2
Probation Reports	2
Case Records	2
Case Conference Reports	2
CYA Policy Manuals	2
Composite Field File	1
Previous Institution Adjustment	1
School Reports	1
Educational Summary	1
Team Meetings	1
Investigations	1
Family History	1
Parole Violation Reports	1
I-Level	1
Available Resources	1
Counseling Reports	1

institution; ward self-evaluation such as Jesness; diagnostic evaluations of probability and prognosis of behavior; interviews with parents; and a summary of all information.

Parole Field Staff

Of the ten parole field staff members interviewed, only four completed the revised format questionnaire. All four rated the Burgess as most easily understood and followed. Two felt that the Burgess allows for easy calculation of a ward's probable parole outcome, one felt that both the Burgess and the predictive attribute permitted such calculations, and one remarked that multiple regression was better in this regard. Two agreed that the use of information dimensions facilitates the comparison of prediction results with subjective hunches, while two disagreed. Three agreed and one disagreed that the generation of score categories for each information dimension allows a comparison of their predictive ability.

Seven of the ten respondents to the professional questionnaire claimed to be aware of statistical prediction research in corrections. Six felt that such research might some day be helpful to the field, three did not, and one was uncertain. General attitudes toward the DPPI concept and formats were positive in three cases, negative in three, and mixed or uncertain in four. Nearly all stated that simplicity was the most important consideration in selecting a format.

The orientation of the majority of parole field staff appeared to be intuitive: two of the ten expressed a preference for objective information sources in their decision

making, while eight indicated that they relied primarily on intuitive or subjective factors. It was generally felt that the kinds of objective information available to them were not particularly helpful to making decisions or evaluating a ward's current situation and that rather than rely on objective information it was more useful to develop a "feel" for clients' current strengths and weaknesses. Five of the respondents stated that if they were to use the index it would be primarily for programming. Only one respondent felt that the self-fulfilling prophecy could be a significant problem with the use of this instrument.

Some of the additional comments were: that too much emphasis was given to case-history factors, which were not as important as offense or other factors, since all offenders had "messed up" backgrounds (one respondent); that the outcome criterion should be arrest rather than commitment, since returns to institutions reflect simply the workings of the judicial process (one respondent); that the index relied too heavily on institutional factors which were not relevant to parole or to the life situation of the ward once returned to the community (four respondents); that the test scores and evaluations contained in the index were not very useful because they were culturally biased (one respondent); that there was not enough emphasis on social/environmental factors, which were of most importance to parole (eight respondents);

that a longitudinal profile reflecting change over time in a ward's score was needed (one respondent); that a reference manual was needed (one respondent); that since the California Youth Authority was shifting its emphasis from rehabilitation to control the instrument should not be geared toward rehabilitative case work (three respondents); that each dimension should have cut-off scores (upper and lower) rather than merely a baseline score (one respondent); and that the ward and all line staff should have access to the DPPI material (six respondents). Table 15 presents, by frequency of mention, information on the types of information used by parole staff in decision making.

Asked what kinds of information not presently available they would like to see included in a ward's file, respondents suggested: more information on ward perceptions and attitudes; more information on family relationships; more emphasis on social/environmental factors; more valid personal information on wards; observations of significant others; more extensive psychiatric workups and intelligence testing; and information on a ward's ability to set, work toward, and reach goals. One respondent stated that what was needed was not more information but more time to evaluate the large amount of information already available. Another noted that what was needed was increased reliability of existing information.

Table 15
Information Used by Parole Field Staff

Type of Information	Frequency of Mention
Interview with ward	10
Arrest Report	5
Clinic Summary	4
Offense History and Pattern	3
Type of Current Offense	2
Academic Achievement	2
Case Files	2
Board Reports	2
Probation Reports	2
Psychological Evaluation	2
Social Worker Report	2
Family Situation	2
Interviews with Significant Others	2
Case Conference Reports	1
Institution Reports	1
Ward Behavior Patterns	1
Ward Progress in Treatment	1
Maturity Level	1
Age	1
Reading Level	1
Drug Use	1
Nonverbal I.Q.	1
Vocational Training or Work	1
Placement Information	1
GATB	1
YA Policy	1
Available resources	1

Reception Center Staff

Six reception center staff were interviewed. None of these completed the format questionnaire, although all responded to the oral questions and the professional questionnaire. Four of the reception center staff were aware of statistical prediction research in corrections; all four believed it could be helpful, although two felt that prediction was not as valid for the youthful ward as for older offenders. The staff members were generally interested in the DPPI project and its concepts but emphasized that any instrument developed would be simply one more tool that might be incorporated into their assessments of the ward. The lack of social context was considered a problem by four of these respondents and five stressed the difficulty of prediction with an age group as young as that with which they were working.

Preference was expressed for the Burgess because of its relative simplicity but the respondents felt that because of the clerical work involved the computation of scores should be centralized and the information perhaps computerized. The types of information used by reception center staff are listed in Table 16.

The respondents indicated that they would like to see the following incorporated into the ward's file: more complete school data (three respondents); initial home visit

Table 16
Information Used by Reception Center Staff

Type of Information	Frequency of Mention
Case File	6
Interview with ward	6
Test Data	5
Prior Offense	3
School Records	3
Behavior Reports	2
Interviews with Significant Others	1
Probation/Parole Reports	1
Psychological Evaluations	1
Institution Reports	1
Interpersonal Relations	1
Staff Observations	1
I.Q.	1
I-Level	1
Home Situation	1
Vocational History	1

report at time of reception; more information on ward's values and orientation; information on vocational history and pattern; more on family and peer relationships (two respondents); information on neighborhood culture; results of earlier contacts with therapists; and Jesness Inventory profile scores.

Parole Board Members

Four parole board members were interviewed as a group to obtain their initial reactions to the DPPI concept and

to determine their interest in the development of such an instrument. All were generally supportive of the effort to develop a clinical-synthesis model and felt that the dimensional approach was more meaningful than a single score such as that provided by the B.E. (which they did not use). They did not feel, however, that parole board members needed such an index at their level of decision making.

ii. Conclusions

As a preliminary step toward development of a clinically relevant predictive device, the DPPI project sought to obtain the practitioner's contribution to what was intended to be a cooperative effort. The field evaluation showed that although many practitioners could see the need for a predictive instrument based on information relevant to their needs, few if any had an interest in the process of developing such an instrument. A common opinion was that someone else should create the instrument (and even score the results), which then might be used by the case worker if it had demonstrated validity and utility.

The one partial exception to this opinion was found among reception center staff, who commented that since the index would be administered at that level they would have to participate in its creation. However, even reception center staff did not want to look at the formats presented

to them or consider them in any detail. Thus, feedback at all levels was restricted to general comments on the utility of parole prediction or operational issues involved in its introduction and use in the field.

The failure to obtain detailed responses on methodological issues apparently was due at least in part to the fact that the practitioner feels he is already overworked and has little time for answering questionnaires or helping researchers do "their" job. A possible solution to this problem might be to provide the practitioner with an incentive to participate in such a cooperative undertaking. The remote incentive of the development of an instrument useful to his work was not sufficient.

An additional barrier to full cooperation was the frequently expressed negative attitude toward research itself, apparently the result of a gap in experience and orientation between the researcher and the practitioner in the field of corrections. Several respondents suggested that the development and testing of such an instrument would have to be undertaken within the correctional setting for the practitioner to be persuaded of its utility.

Despite the disappointing results on the comparison of the alternate formats and on the consideration of the design of a model instrument, field staff made numerous interesting suggestions and observations which may be helpful in further

efforts to create a clinically useful prediction instrument. The most obvious example, because of its almost unanimous support at all levels, was that any instrument to be used by the clinician must be simple in format. Simplicity was seen to be more important than any consideration of how the dimensions were derived or how the scores were obtained. Almost all of the respondents who expressed a preference for one of the three suggested formats selected the Burgess format for its relative simplicity, although the even more simple graph was generally preferred. It was frequently stated that even the Burgess involved too much clerical work for line staff and that "centralization," and perhaps even computerization, of the data required for deriving scores would be necessary.

All levels of staff interviewed expressed some interest in the definition of the outcome criterion. Parole field staff tended to view re-arrest as the criterion of interest, while institutional staff focused on revocation of parole and reception center staff recommended consideration of a continuum of outcomes rather than a simple dichotomy based on arrest or violation. Although this project assumed that parole outcome was of at least some concern at all points of the correctional sequence, interview responses suggested that field staff considered other factors to be at least as important or more so. Many respondents, and especially those

at the reception center, felt that they were too far removed from the point of parole success or failure to influence parole outcome. Reception center and institutional staff appeared to be somewhat frustrated by the lack of feedback from later stages of correction concerning the progress and ultimate outcome of wards with whom they had contact. Field staff at each level indicated that the DPPI model should incorporate some provision for feedback on outcome.

Another issue noted by many respondents, and recognized in advance by project staff, concerned the data base utilized in the DPPI project. While it is an unusually rich data base with considerable potential for varying analyses, the kinds of information available when this data base was developed (1964-65) were no longer being collected by the California Youth Authority at the time of the DPPI study. For the purposes of the creation of a model for use in later development of an instrument, the data base was considered most appropriate. However, the observation by field staff that the data would have to be updated before the index would be usable was correct.

Almost all of the respondents who examined the proposed model approved the dimensional approach but noted the scarcity of the type of information with which they were greatly concerned: the individual's social environment and his interactions within it. All respondents' comments on the

kinds of presently available information they would like to see incorporated into the ward's file must therefore be considered and the information integrated into any index developed for use in the field.

The consideration of the youthfulness of the population with which these field staff worked was stressed as important in the design of an index to predict parole outcome. The Base Expectancy has been rated as less appropriate for use with young populations than with older ones because of the greater likelihood that young wards will change during correctional treatment. Field staff at all levels noted the need for longitudinal profiles reflecting change during treatment or for dimensional items based on dynamic factors which may change from one administration of the instrument to another. While the importance of these considerations was recognized by project staff, because of the nature of the existing data base few such items could be included. Future efforts which involve the collection of new data should consider the incorporation of information on change factors to reflect progress in treatment or the effects of maturation.

VI. A CRITICAL ASSESSMENT OF THE DPPI METHODOLOGY

This project set several ambitious objectives and, as is often the case with research of this scope, did not achieve all

of them. When research objectives go unmet it is often fruitful to examine the methodology of that research in an attempt to discover whether other methods might yield the desired results. The attempt here is to use the experience of this project to identify obstacles and to suggest ways in which they might be negotiated. The objectives of this project were to:

1) demonstrate that predictive information can be fused with clinical concerns using a clinical synthesis model;
2) compare the accuracy of several predictive methods and to examine the trade-off in predictive power which results from restricting predictors to clinically relevant dimensions;
3) design and test several formats for displaying predictive information in a readily understandable and useful manner.

Objective 1

The first objective involved several parts: (a) the development of the theoretical basis for clinical synthesis; (b) the development of a rationale for the application of a clinical synthesis model in correctional case decisions; (c) explication of the operation of the model in the correctional setting; and (d) a field test of the theoretical basis, rationale, and application of the DPPI, leading to acceptance,

modification, or rejection of them and their constituent elements.

Steps a, b, and, to a lesser extent, c were accomplished. The foregoing presented the theoretical basis for clinical synthesis and the rationale for its application to correctional case study issues, and described potential operation in the field.

The field test, which was intended to allow complete specification of field operation and to provide data for evaluating the theoretical basis and rationale, was not completely successful. Correctional field personnel, on whom the study design relied for real-world data needed for these tasks, were unable or unwilling to devote the time required to fully meet study objectives. As a result the DPPI's theoretical basis and rationale remain essentially untested.

The data collected indicate that substantial refinement of operational methods will be required for successful demonstration that predictive information can be fused, through a clinical synthesis model, with case study concerns. The refinement of operational procedures and thorough field testing of the theory and rationale developed above will require a commitment of substantial time and energy on the part of operational personnel.

A major methodological shortcoming of this study was that it did not elicit adequate cooperation from correctional staff.

In retrospect it is clear that the assumption that line personnel could absorb the additional workload of detailed, repetitive review of the DPPI formats, self-interrogation regarding how they make decisions, and identification of the types of information they use or would like to use in decision making was unrealistically optimistic.

A more useful methodology would have recognized the time constraints under which operational staff function. The study design could then have ensured the required input from field staff by compensating them. This could be done either by hiring operational personnel as consultants or by providing funding to correctional agencies which would allow the reduction of regular workload for staff participating in the research effort.

Objective 2

Like the first, this objective was comprised of several subparts: (a) the development of predictions from seven dimensions presumed to have clinical relevance; (b) the development of predictions without regard to clinical relevance of predictor variables; (c) the comparison of predictions from the seven dimensions with the overall predictions in order to estimate the loss in predictive power resulting from restricting predictors to clinically relevant dimensions; (d) the comparison of the power of several prediction methods; and (e) the evaluation of the hypothesis of "masked heterogeneity."

Each of these subtasks was accomplished with a substantial degree of methodological rigor. The value of their accomplishment is limited, however, because of the low predictive power which resulted from the various predictive efforts. This low power, considerably lower than found in numerous other studies, makes it unwise to accept the results obtained here as conclusive.

Since the study methodology made no provision for this unanticipated outcome, the issues which the study sought to address remain unresolved. This outcome can be attributed to four basic methodological decisions. First, the predictive power of the data used in this study was not pretested. Second, the study design was ex post facto and it was thus not possible to expand the set of potential predictors when the original variables proved to be of limited value. Third, the data used were on youthful offenders, whose behavior has been found to be less predictable than that of older persons. Finally, the parole follow-up consisted of a relatively short fifteen-month period, which in future studies should be substantially increased. These criticisms should be evaluated with the realization that the use of alternative data would have been much more costly since the data used in this study were collected, coded, and in machine-readable form from the start. In addition, this data set is rich in various test scores which were

thought to enhance the value of the data set for an effort such as this where clinical relevance and appeal to case workers were important considerations.

Use of a prospective design would have allowed for the expansion of the data but would have required a much longer study at greater cost. In retrospect, such an expanded study would appear to be required for maximum flexibility.

Objective 3

The least satisfactory of all the efforts undertaken in the study was the field testing of alternative formats for the presentation of predictive information. It had been planned that through interaction of field staff and research staff much progress would be made in efforts to develop a synthesis of prediction research and case-study practice. What occurred was a superficial, hurried review by operational personnel of the various DPPI formats. This failure originated in the study design, which had not reckoned with the time demands or skepticism of field staff. Clearly, the methodology was flawed in that it did not encompass mechanisms to relieve the time pressures or to demonstrate the need, value, and utility of fusing prediction and case-study concerns.

Conclusion

The foregoing discussion of the objectives of the DPPI

and the extent to which they were achieved makes apparent two shortcomings of the study methodology which must be avoided in future attempts to build a synthesis of prediction and case-study concerns. These are the failure to include correctional staff early in the planning phase as well as to ensure their full, on-going participation in the study, and the decision to use a fixed set of data.

The lack of adequate participation by field staff leaves the theoretical basis and rationale untested and precludes satisfactory exploration and refinement of the DPPI formats.

The fixed data base used here proved to be unsatisfactory because it precluded the possibility of using additional data items to improve predictive power. Without the opportunity to explore this possibility it cannot be conclusively determined whether parole outcome is unpredictable for this group or merely depends on other factors.

Future studies could rectify these methodological inadequacies. Such future studies would be particularly valuable if they included a longer parole follow-up period, which could shed new light on predictive issues as well as provide important knowledge about criminal careers.

VII. THE FUTURE OF PAROLE PREDICTION

A. Relative Usefulness of Clinical and Actuarial Parole Prediction Techniques

Attempts to develop actuarial prediction devices appropriate for decisions considered to be within the province of case workers often generate discussions of the relative value of clinical and actuarial methods of prediction. It is our view that such debate is misdirected since it addresses a question not properly at issue, i.e. the question of which approach is superior. This question supposes that the two are basically dissimilar, a supposition rejected above. It implies also that the superior method should be used exclusively. Such a conclusion is unrealistic since the ethical foundations of western society require the insertion of a human decision maker when formal decisions regarding individual freedom are made.

It seems clear that efforts are better made to integrate the two approaches than to place them in opposition to one another. Both the actuarial and clinical approaches can provide valuable input to correctional decisions. Much prior research has shown that actuarial predictions are more accurate than unaided clinical judgment. For this reason it would seem a dereliction to omit actuarial input. It is equally clear that the clinical role is also needed because the human decision maker contributes the ability to respond to unique situations, an appreciation of multiple criteria, the possibility of therapeutic input, and a humanizing influence on decisions.

For these reasons we do not consider the question to be

one of clinical or actuarial prediction but rather of how best to combine the two approaches so that the unique value of each is preserved. There is substantial reason to believe that such a combination is possible, that it can improve decisions, and that fusion of the two might prove synergistic.

B. The Utility of Further Efforts to Develop a Combined Prediction Technique

In light of the mixed outcomes of this project, one might suppose that we would question the value of further efforts in this vein. This supposition would be far from correct. Despite the disappointments of this effort, we remain firmly convinced that such efforts should continue for the following reasons:

1. The superiority of actuarial prediction of structured criteria from historical data is well established.

2. The correctional clinical setting is characterized by a limited number of highly structured decisions and a much greater number of unstructured, situational decisions.

3. Human decision makers are limited in their ability to recall and analyze the vast amount of data necessary to discover and quantify complex relationships in the data.

4. Actuarial approaches are often not adequate to recognize atypical situations or to quickly respond to changes in underlying processes.

5. Research in artificial intelligence (one part of

which attempts to develop self-correcting computer systems which improve with experience) has suggested that mechanistic systems are far superior to humans in highly deterministic situations where outcomes can be exhaustingly analyzed, but that a non-mechanistic combination is better able to respond to dynamic, multicriterion, highly variable decision settings.

These considerations, we believe, strongly argue that we continue to search for a practical synthesis of actuarial and clinical approaches.

Considerable evidence exists that such syntheses are possible--the Parole Decision-Making Project (Gottfredson et al., 1973) demonstrated that actuarial techniques could be used to make previously implicit policy explicit, guide case decisions, and provide a structure in which past decisions and their consequences be better understood.

Work done in conjunction with the Virginia Department of Corrections also has demonstrated the potential value of using an actuarial device in correctional programming (Brookhart et al., 1976).

C. Areas of Future Research and Methodological Approaches

The strongest lesson of the DPPI is that attempts to fuse case-work and actuarial approaches must include clinical personnel as full partners in the effort. Such attempts must appreciate the context of day-to-day case work and must provide practical benefits in this context. A forceful demonstration of this was the preference of operational personnel

for the simplest possible DPPI formats because other formats were viewed as too complex and too time consuming to be practical. This response, rather than reflecting lack of interest on the part of the respondents, clearly reflects operational realities.

This all suggests that the first order of business is further research into clinical decisions, process, and needs. This research should consider such issues as:

1) the operational, day-to-day decision needs of case workers;
2) the various criteria which are important in different settings;
3) the structure for providing information which is most useful for various decisions;
4) the types of theories and constructs used by case workers and how actuarial approaches can support and refine their use;
5) how to motivate the use of new tools and techniques in the case study setting (reduction of workload through computerization of information retrieved and index preparation);
6) how to provide feedback about prior decisions in a non-threatening manner which encourages improved decisions.

When the foregoing issues have been addressed adequately the following areas seem worthy of investigation:

1) the applicability of the clinical synthesis model;
2) whether provision of DPPI-type information does support the development of productive stimulus-response chains in the interview setting;
3) whether case workers can unearth case-specific factors which should mitigate actuarial predictions;
4) matching of case worker and client on the basis of prior successful relationships;
5) clinically directed predictions in which the case worker would identify the variable of particular balance for a given case; these variables would then be employed in developing an actuarial prediction;
6) predictors based on variables identified by theory;
7) inclusion of change and "progress" variables in predictions;
8) identification of homogeneous subgroups with unique predictive relationships;
9) the relationship of clinical approach and preference for different information presented in different fashions;
10) the consideration of long parole follow-up periods and research into variables that influence criminal

career patterns;

11) evaluation of criminological theory.

These issues must be central to any future attempt to fuse clinical and case-study concerns. Research undertaken without considering these issues and without learning how to use prediction information might yield only small increases in predictive accuracy and therefore be limited to academic interest.

REFERENCES

Allport, G. W. "The psychologist's frame of reference." Psychology Bulletin 37: 1-28, 1940.

Allport, G. W. Pattern and growth in personality. New York: Holt, Rinehart & Winston, 1961.

Babst, D. V., Gottfredson, D. M., & Ballard, K. B. "Comparison of multiple regression and configural analysis techniques for developing base expectancy tables." Journal of Research in Crime and Delinquency, 5(1): 1968.

Benson, Sir George. "Prediction methods and young prisoners." British Journal of Delinquency, 9(3): 1959.

Beverly, R. F. Base expectancies and the initial home visit schedule. Sacramento: California Youth Authority, 1964.

Beverly, R. F. & Guttmann, E. S. An analysis of parole performance by institution of release. Sacramento: California Youth Authority, 1962.

Brookhart, D. E., Ruark, J. B., and Scoven, D. E. A strategy for the prediction of work release success, final report. Va. Dept. of Corrections, 1976.

Burgess, E. W., in Bruce, Harno, Landesco, & Burgess. The workings of the indeterminate sentence law and the parole system in Illinois. Springfield, Ill.: Board of Parole, 1928.

Chein, I. "The logic of prediction: some observations on Dr. Sarbin's exposition." Psychology Review, 52: 1945.

Cockett, R. "Borstal training: a follow-up study." British Journal of Criminology, 7(2): 1967.

Coyle, E. Counselor prediction of academic success. New York: Columbia University Ph.D. dissertation, 1956.

Dean, C. W. "New directions for parole prediction research." Journal of Criminal Law, Criminology and Police Science, 59(2): 1968.

DeGroot, A. D. "Statistical via clinical prediction." Invited address, Western Psychological Association, San Jose, California, April 1960.

Duncan, O. D., and Duncan, B. "A methodological analysis of

segregation indexes." American Sociological Review 20: 210-217, 1955.

Duncan, O. D., Ohlin, L. E., Reiss, A. J., Jr., and Stanton, H. R. "Formal devices for making selection decisions." American Journal of Sociology 48: 573-584, 1953.

Evjen, V. H. "Current thinking on parole prediction tables." Crime and Delinquency, 8(3): 1962.

Gibbons, T. C. N. "Prediction studies and psychiatric diagnosis." In Grygier, T. et al., eds. Criminology in transition. London: Tavistock, 1965.

Glaser, D. "A reconsideration of some parole prediction factors." American Sociological Review, 19: 1954.

Glaser, D. "The efficacy of alternative approaches to parole prediction." American Sociological Review, 20: 1955.

Glaser, D. "Differential association and criminological prediction." Social Problems, 8: 1960.

Glaser, D. "Prediction tables as accounting devices for judges and parole boards." Crime and Delinquency, 8(3): 1962.

Glaser, D. The effectiveness of a prison and parole system. New York: Bobbs-Merrill, 1964.

Glueck, S. & Glueck, E. 500 criminal careers. New York: Alfred A. Knopf, 1930.

Glueck, S. & Glueck, E. Unraveling juvenile delinquency. New York: Commonwealth Fund, 1950.

Gottfredson, D. M. "Assessment and prediction methods in crime and delinquency." In U.S. President's Commission on Law Enforcement and Administration of Justice. Task Force Report: Juvenile Delinquency and Youth Crime. Wash., D.C.: U.S. Government Printing Office, 1967.

Gottfredson, D. M. & Ballard, K. B. The validity of two parole prediction scales: an eight year follow-up study. Institute for the Study of Crime and Delinquency, 1965.

Gottfredson, D. M., Ballard, K. B., and Lane, L. Association analysis in a prison sample and prediction of parole performance. Institute for the Study of Crime and Delinquency, 1963.

Gottfredson, D. M. and Beverly, R. F. "Development and operational use of prediction methods in correctional work." Proceedings of the Social Statistics Section. Wash., D.C.: American Statistical Association, 1962.

Gottfredson et al. Parole decision-making. National Council on Crime and Delinquency, Davis, California, 1973.

Gough, H. G. "Clinical versus statistical prediction in psychology." In Postman, L. (ed.), Psychology in the making. New York: Knopf, 1962.

Gough, H. G., Wenk, E. A., and Rozynko, V. V. "Parole outcome as predicted from the C.P.I., the M.M.P.I., and a base expectancy table." Journal of Abnormal Psychology, 70(6): 1965.

Hart H. "Predicting parole success." Journal of Criminal Law and Criminology, November 1923.

Helmstadter, G. C. "An empirical comparison of methods for estimating profile similarity." Educational and Psychological Measurement, 17: 1957.

Hoffman, P. J. "Criteria of human judgment ability: (1) the clinical assessment of intelligence and personality." American Psychologist, 13: 1958.

Hoffman, P. J. "The prediction of clinical prediction." American Psychologist, 14: 1959.

Holt, R. R. "Clinical and statistical prediction: a reformulation and some new data." Journal of Abnormal and Social Psychology 56: 1-12, 1958.

Holt, R. R. "Yet another look at clinical and statistical prediction: Or, is clinical psychology worthwhile?" American Psychologist 25(4): 337-349, 1970.

Hood, R. Homeless borstal boys. London: Bell, 1966.

Hutt, M. L. "Actuarial and clinical approaches to psychodiagnosis." Psychological Rep. 2: 413-419, 1956.

Klein, G. S. "An application of the multiple regression principle of clinical prediction." Journal of General Psychology, 38: 1948.

Klopfer, B., Kirchner, F. J., Wisham, W., and Baker, G. "Rorschach prognostic rating scale." Journal Proj.

Technology 15: 425-428, 1951.

Laune, F. F. Predicting criminality. Evanston, Ill.: Northwestern University, 1936.

Lejins, P. P. "Parole prediction: an introductory statement." Crime and Delinquency, 8(3): 1962.

MacNaughton-Smith, P. "The classification of individuals by the possession of attributes associated with a criterion." Biometrics, June 1963.

MacNaughton-Smith, P. Some statistical and other numberical techniques for classifying individuals. London: Her Majesty's Stationery Office, 1965.

Mannheim, H. and Wilkins, L. T. Prediction methods in relation to borstal training. London: Her Majesty's Stationery Office, 1955.

Meehl, P. E. Clinical vs. statistical prediction: a theoretical analysis and a review of the evidence. Minneapolis: University of Minnesota Press, 1954.

Meehl, P. E., and Dahlstrom, W. G. "Objective configural rules for discriminating psychotic from neurotic MMPI profiles." Journal of Consulting Psychology 24: 375-387, 1960.

McArthur, C. C. "Comment on studies of clinical versus statistical prediction." Journal of Counseling Psychology 15 (2): 172-173, 1968.

Monachesi, E. D. Prediction factors in probation. Hanover, N.H.: The Sociological Press, 1932.

Mosel, J. N., and Roberts, J. B. "The comparability of measures of profile similarity: an empirical study." Journal of Consulting Psychology, 18: 1954.

Muldoon, J. F., and Ray, O. S. "A comparison of pattern similarity as measured by six statistical techniques and eleven clinicians." Educational and Psychological Measurement, 18: 1958.

Ohlin, L. E. Selection for parole. New York: Russell Sage Foundation, 1951.

Ohlin, L. E., and Duncan, O. D. "The efficiency of prediction in criminology." American Journal of Sociology, March 1949.

Ohlin, L. E., and Lawrence, R. A. "A comparison of alternate methods of parole prediction." American Sociological Review, 17(3): 1952.

Pankoff, L. D., and Roberts, H. V. "Bayesian synthesis of clinical and statistical prediction." Psychological Bulletin 70:6 762-773, 1968.

Polyani, M. Personal knowledge: towards a post-critical philosophy. (Rev. ed.) New York: Harper Torchbooks, 1964.

Powers, S. B. "Standard parole prediction methods: views of a correctional administrator." Crime and Delinquency, 8(3): 1962.

Reiss, A. J. "Unraveling juvenile delinquency II: an appraisal of research methods." American Journal of Sociology, 57: 1951.

Rogers, J. W. "Parole prediction in three dimensions: theory, prediction, perception." Sociology and Social Research, 52: 1967.

Sanders, B. S. "Testing parole predictions." Proceedings of the 65th Annual Congress of the American Prison Association, 1935.

Sarbin, T. R. "Clinical psychology--art or science?" Psychometrika 6: 391-400, 1941.

Sarbin, T. R. "A contribution to the study of actuarial and individual methods of prediction." American Journal of Sociology, 48: 1942.

Sarbin, T. R. "The logic of prediction in psychology." Psychological Review 51: 210-228, 1944.

Sarbin, T. R., and Taft, R. An Essay on Inference in the Psychological Sciences. Berkeley, California: Garden Library Press, 1952.

Sawyer, J. "Measurement and prediction, clinical and statistical." Psychological Bulletin 66:3 178-200, 1966.

Schuessler, K. F. "Parole prediction: its history and status." Journal of Criminal Law and Criminology, November 1954.

Simon, F. H. Prediction methods in criminology. (Home Office Research Unit Report). London: Her Majesty's Stationery Office, 1971.

Stouffer, S. A. "Notes on the case study and the unique case." Sociometry, 4: 349-357, 1941.

Thompson, R. E. "A validation of the Glueck Social Prediction Scale for proneness to delinquency." Journal of Criminal Law, Criminology and Police Science 43(4): 451-470, 1952.

Tibbetts, C. "Success or failure on parole can be predicted." Journal of Criminal Law and Criminology, May 1931.

Trankell, A. "Psychologist as an instrument of prediction." Journal of Applied Psychology 43: 170-175, June 1959.

Vold, G. B. Prediction methods and parole. Hanover, N.H.: The Sociological Press, 1931.

Vold, G. B. "Do parole prediction tables work in practice?" In Publication of the American Sociological Society, 1931.

Voss, H. L. "The predictive efficiency of the Glueck Social Prediction Scale." Journal of Criminal Law, Criminology and Police Science, 54(4): 1963.

Wallin, P. "The prediction of individual behavior from case studies." In P. Horst et al., The Prediction of Personal Adjustment. New York: Social Science Research Council Bulletin No. 48, 1941, pp. 183-239.

Ward, P. G. "The comparative efficiency of differing techniques of prediction scaling." Australian and New Zealand Journal of Criminology, 1(2): 1968.

Warner, S. B. "Factors determining parole from the Massachusetts Reformatory." Journal of Criminal Law and Criminology, August 1923.

Webster, H. "Rao's multiple discriminant technique applied to three TAT variables." Journal of Abnormal Social Psychology, 47: 1952.

Wilkins, L. T. Evaluation of penal measures. New York: Random House, 1969.

Wilkins, L. T., and MacNaughton-Smith, P. "New prediction and classification methods in criminology." Journal of Research in Crime and Delinquency, 1: 1954.

Winkler, R. "The assessment of prior distributions in Bayesian analysis." Journal of the American Statistical Association 62: 776-880, 1967.

Wittman, M. P. "A scale for measuring prognosis in schizophrenic patients." Elgin Papers 20-33, 1941.

Wittman, M. P., and Steinberg, L. "Follow-up of an objective evaluation of prognosis in dementia praecix and manic-repressive psychoses." Elgin Papers 5: 216-227, 1944.

Zubin, J. "Clinical vs. actuarial prediction: a pseudo problem." In Proceedings of 1955 Invitational Conference on Testing Problems. Princeton: New Jersey, Educational Testing Service, pp. 107-128, 1956.