

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/263191421>

Developing an Empirically Based Typology of Delinquent Youths

Article in *Journal of Quantitative Criminology* · September 1999

DOI: 10.1023/A:1007524414432

CITATIONS

32

READS

2,740

2 authors:



Peter Jones

Temple University

71 PUBLICATIONS 525 CITATIONS

[SEE PROFILE](#)



Philip W. Harris

Temple University

38 PUBLICATIONS 649 CITATIONS

[SEE PROFILE](#)

Developing an Empirically Based Typology of Delinquent Youths

Peter R. Jones¹ and Philip W. Harris¹

A number of personality-based typologies have been developed to differentiate juvenile delinquents. Some of these typologies have been faulted for their method of construction, for their failure to provide outcome measures, or because they are dated and have not been retested with more recent delinquent populations. This paper describes efforts to develop an empirically derived, inductive typology of juvenile delinquents that can be used in applied settings to facilitate a better understanding of program outcomes and enable placement authorities better to predict program effectiveness. The data employed in the analysis come from an outcome-based information system—*ProDES*—that has been operational for more than 4 years in Philadelphia. Unlike many such databases, *ProDES* was jointly designed by researchers and practitioners to serve a large system of community and institutional programs. Included among the outcome measures are four standardized, self-reported personality scales that comprise 14 subscales and form the basis of the typology. This paper describes the methods used to construct the typology, the substance of the classification system, and its potential utility for program evaluation.

KEY WORDS: juvenile delinquency; empirically based typology; outcome-based information system; cluster analysis.

1. INTRODUCTION

Classification is central to the behavioral sciences and, whether we recognize it or not, it is fundamental to juvenile justice (D. M. Gottfredson, 1987; Brennan, 1987; Everitt, 1993). The specific purpose of classification varies enormously, ranging from the testing of delinquency theories (Megargee *et al.*, 1979; Warren and Hindelang, 1979; Jefferson and Johnson, 1991; Huizinga *et al.*, 1991) and improvement in the matching of offenders to interventions (Sechrest, 1987; Harris, 1988; Brannon *et al.*, 1989; Mezzich *et al.*, 1991; Van Voorhis, 1997; Palmer, 1992, 1984) to improvement in offender population management (Baird, 1986; Glaser,

¹Department of Criminal Justice, Gladfelter Hall, Temple University, Broad & Montgomery, Philadelphia, Pennsylvania 19122.

1987; Dembo *et al.*, 1994) and better prediction of delinquent behavior (Warren, 1965; D. M. Gottfredson, 1987; S. D. Gottfredson, 1987; Brennan, 1987; MacKenzie *et al.*, 1988; Andrews *et al.*, 1990; Palmer, 1992; Jones, 1996). Unfortunately, despite these complex and multiple goals, the classifications actually developed and used in juvenile justice involve little more than “judgmental” approaches that specify a level of repetitive offending or the identification of an offender type such as sex offender or drug abuser (Clements, 1996). Using these basic categorizations, decision-makers such as judges, program directors, and system administrators must decide among a broad range of interventions.

Gottfredson (1987) has argued that advances in classification mark the milestones in the history of research in criminology and criminal justice. From the early beginnings of Goring (1913) and Lombroso (1911), our interest in classification has developed to include the effectiveness of differential treatment for particular categories of offenders (Adams, 1961; Warren, 1971). The latter body of work is based upon the premise that certain interventions work for certain persons under certain conditions. Indeed, the question of what works with whom in what circumstances (Palmer, 1992; Van Voorhis, 1994; Andrews *et al.*, 1990) remains one of the most critical elements of program evaluation in juvenile corrections and delinquency prevention.

Much of the work on classification tends to adopt one of two basic approaches. First, there are the classifications that are developed from theoretical conceptualizations about the precursors to delinquency or of the social and personal factors that impact treatment outcomes. Second, there are the classifications that are empirically derived, often more atheoretical and with a specific objective (usually prediction of delinquency) in mind. Despite the voluminous literature weighing the advantages of one approach or one classification system against another, the fact remains that there is no single “best” classification system. As Brennan (1987) has argued, the goal is to select the method that best fits the research data and purpose.

The purpose of the present study is to address a specific problem identified by Palmer (1992) and others in their work on program intervention and effectiveness. Unfortunately, most program designs as well as most research evaluations fail to differentiate among program clients, thus masking important differential outcomes (Palmer, 1992). Programs are therefore judged to be effective or ineffective without taking into account the likely interactions between type of client and type of program that may well occur. As Andrews (1990) has noted, we continually produce findings that appear to support the “nothing works” conclusion and that miss the fact that some things work, for some people, some of the time.

Even if one believes that classification can help improve our ability to evaluate programs we are still left with the question—What type of classification? There exist already a large array of juvenile classification systems, though many such as the I-level (Warren, 1976) and the juvenile typology (Cloward and Ohlin, 1960) date from the “subjective assignment” phase of the pre-1980s and have been heavily criticized for poor reliability and validity (Becker and Hyman, 1972). Our goal, therefore, is to construct empirically a treatment-oriented typology that would provide an alternative approach to classifications based on theoretical speculation.

In this article, we describe our approach and our methods in developing the typology and we present some results from its implementation in the field.

2. CLASSIFICATION FOR TREATMENT—RECENT ISSUES

Though there is widespread agreement that classification can improve the overall effectiveness of interventions, it is generally agreed that little development in this area has occurred in the past two decades (Van Voorhis, 1997; Clements, 1996; MacKenzie *et al.*, 1988; Sechrest, 1987). Early developments in multivariate taxometric methods were generally directed toward predictive classifications (e.g., Wilkins and MacNaughton-Smith, 1964) and consequently risk assessment has tended to capture both the imagination and the research funds of Federal and State agencies. Far less developmental work has been conducted in the area of treatment-oriented classification. Of course, given the methodological criticisms and overall skepticism leveled at existing treatment-oriented taxonomies, it comes as no surprise that so little developmental work exists. At present, the Jesness Inventory and the MMPI represent two of the most commonly used classification systems used in juvenile justice, a fact not unrelated to their reputation as objective, standardized measures with known levels of validity and reliability.

3. METHODS OF TYPOLOGY DEVELOPMENT

Evaluations of classification systems tend to focus on two issues—the extent to which they are grounded in criminological theory and the extent to which they adopt an appropriate methodological approach to the task at hand.

Sechrest (1987) has noted that classification systems have generally developed without any theoretical link to interventions and he suggests not only that should a classification system be grounded in criminological theory but that programs should be developed to target the specific needs identified by the classification system. Such an approach to treatment is

supported by Clements (1996), who adds that knowledge of what has transpired inside the treatment experience is also necessary in order to assess the interaction between person type and treatment method. One recent classification system to incorporate this approach is the Level of Supervision Inventory (LSI) of Andrews and Bonta (1994), a risk and needs typology that is specifically relevant to social learning and cognitive approaches.

From a methodological perspective we can identify four crucial phases in the development of a typology—the need to identify relevant measures, to utilize appropriate techniques, to test for internal validity, and to test for external validity (Brennan, 1987). We deal with each of these phases in detail as we describe the development of the present typology.

4. THE CONTEXT OF THE RESEARCH

The typology we present was developed from data derived from the *ProDES* (Program Development and Evaluation System) research project in Philadelphia (Harris and Jones, 1996). Started in 1991 the project involved the design, implementation, and utilization of a large-scale juvenile justice information system that would enable ongoing monitoring and assessment of all juvenile delinquent programs utilized by Philadelphia's Department of Human Services.² The information system represented a dramatic departure from the more traditional approach of discrete, cross-sectional evaluations of individual programs.

From the outset *ProDES* combined inductive and deductive approaches to evaluation. The content of the information system was developed jointly by the researchers and the programs (with both bringing to bear their particular theoretical and practical orientations). A series of evaluability assessments³ conducted in a sample of programs made it possible to identify a number of information fields on which there was both theoretical and practical consensus. Essentially, program staff were asked to discuss the theoretical underpinnings of their programs and to define those features of both juveniles and programs that they felt were crucial to an understanding of the juvenile justice "business." Their answers shaped both the structure and the content of *ProDES*. The result was a system that comprised measures of personality, risk, and need as well as more common prior record and current offense measures. In addition, the system was designed to measure change in the juveniles by collecting data at three points in time (program entry, discharge, and a 6-month follow-up). Since the typology is

²The population included in *ProDES* comprises all juveniles whose disposition involved commitment to a private or state agency. Juveniles receiving regular probation disposition are not included in the system.

³For a description of this technique, see Wholey (1994) and Smith (1989).

based on measures collected at the point of program intake, we restrict our discussion of the data elements to those utilized in the analysis.

The evaluability assessments confirmed the importance of a number of measures independently identified in the delinquency literature. Without exception, programs told us that self-esteem and values were two key dimensions in their interventions with juveniles. They pointed to school and education dimensions that were centered far more on the beliefs and attitudes of the juveniles than on issues of attendance and performance. Finally, they consistently mentioned measures of family relations, once again emphasizing attitudinal rather than more traditional family composition dimensions.

The measures described as important by the programs resonate with two prominent theoretical perspectives on delinquency. The Family and School Bonding scales were derived from Hirschi's (1969) social control theory, specifically measuring the constructs of attachment, commitment and involvement. The Values Orientation scale taken from the Jesness Inventory is grounded in the theory of delinquent subcultures (Miller, 1958). Consistent with a control theory perspective, however, individuals who score in the antisocial direction on this scale are believed to take little responsibility for their own behavior (Jesness, 1983). Finally, the measure of self-esteem rests on the theoretical proposition that individuals with poor self-image are more vulnerable to the influence of others.

We searched the literature for existing scales that met our specific needs. Our goal was to identify measures that were robust enough for use in an ongoing, large-scale project involving data collection in many sites and with juveniles ranging widely in both age and educational ability. We pilot tested a variety of potential scales finally agreeing on four. For self-esteem we selected the Rosenberg (1965) scale, scoring it as a simple additive rather than a Guttman scale.⁴ The values dimensions was measured using the Values Orientation subscale from the Jesness (1983) Inventory. For attitudes to school we utilized the seven subscales of school bonding developed and reported by Giordano and Cernkovich (1987). Finally, for attitudes to family we used the five subscales of family bonding developed and used by Cernkovich and Giordano (1992). Each scale has been used in *ProDES* for over 4 years and a recent assessment of their performance confirmed their validity⁵ and reliability⁶ (Jones and Harris, 1997).

⁴Validated results from prior adaptations of the scale using a 10-point additive scoring system are reported by Robinson and Schaver (1973).

⁵Factor analyses of all individual items confirmed the factor structure suggested by the scales and demonstrated their statistical as well as conceptual independence. Cross-correlation of the scale scores suggested a high degree of internal validity. Details of the analysis are given by Jones and Harris (1997).

⁶Cronbach's α for the scales varied from a low of 0.60 for family bonding "control and supervision" to a high of 0.85 for the values orientation scale. All α 's indicated acceptably reliable scales.

5. DEVELOPING THE *ProDES* TYPES

Earlier we noted that classification has several purposes, from grouping offenders in terms of numbers of prior offenses to more complex goals of prediction and etiology. Brennan (1987) has noted the importance of selecting the appropriate method of statistical technique for the purpose of the classification. From the wide array of multivariate taxonomic tools available, we selected cluster analysis as the vehicle to the development of our classification.

Everitt (1993) argues that the goal of any numerical method of classification is to provide objective and stable classifications. Objectivity is derived from analyses that study the same set of numerical data with the same sequence of methods and yield the same results. Stability requires that classifications remain the same even after the addition of new cases. The former goal is easily achieved through cluster analysis, but the latter goal is somewhat illusory. When cluster analysis is run on a sample of juveniles, the distribution of scores on the cluster variables will determine the multivariate space within which the classification can occur. As one adds new cases the stability of the classification is possible only if the new cases approximate the same multivariate distributions as the original sample. If, however, the distribution of personality scores changes over time (as is likely), then the addition of new cases will change the multivariate space and, therefore, the classification itself. This has not proven to be an issue in most prior classification research since the analyses have been based upon a discrete sample of cases. Norms have been established and subsequent cases classified according to static category parameters. In the present study we recognize that all classification systems are a product of a particular time and are as susceptible to change as the personalities upon which they are based. One unique feature of an information system such as *ProDES* is its ability to identify and monitor such changes over time.

In the present analysis we used the 14 separate personality measures (self-esteem, values, 7 school bonding subscales, and 5 family bonding subscales) to identify empirically a number of juvenile types. At the outset we did not know how many groups might exist, how each group would be characterized, or how meaningful each group might be. Each of these issues is addressed in the analysis described below.

6. THE CLUSTER ANALYSIS

The cluster analysis was performed with the K-mean algorithm of SPSS 8.0, using the Euclidean distance function and nearest centroid sorting. The

approach involved the following steps.

1. A principal-components analysis (PCA) conducted on the 14 original personality subscales known as the “basis” variables.
2. A cluster analysis of the newly created principal components.
3. A series of diagnostic tests to establish the optimal cluster solution.
4. Fine-tuning of the solution by examination and removal of outliers and respecification of cluster centroids.
5. Definition of the final cluster solution in terms of the 14 basis variables and correlation analysis between the clusters and other case characteristics, such as demographic, family, and offense data.
6. External validation of the classification through the use of a focus group of clinical staff to characterize identified “clusters” based on their familiarity with the youths.

6.1. Principal-Components Analysis

Principal-components analysis (PCA) of the original variables was used to identify redundant or highly correlated variables. The use of highly correlated variables in cluster analysis effectively weights the classification toward a single latent dimension, thereby introducing bias. Further, the use of the original variables would ensure that school bonding (seven subscales) played a far stronger role in shaping the classification than the single measure of values. A third benefit of the PCA is that it standardizes the metrics used for each of the measures (the original scales involve 5-point scales for school and family bonding, a 10-point scale for self-esteem, and a 30-point scale for values).⁷

The PCA with varimax rotation was performed on the original 14 scales and identified four significant components (see Table I).⁸ The first component included four of the five family bonding measures. The second component included the self-esteem and values variables along with two school bonding variables (perceived opportunities and consequences of arrest) and one family bonding variable—identity support. The third and fourth components comprised three and two variables, respectively, from the school bonding scale.

The four significant principal components explained 58% of the original variance in the data. The four component scores were used in all stages of

⁷Clustering algorithms assume interval level measures. Any nominal- or ordinal-level measures should be rescaled to metric measures before cluster analysis begins. Using principal components from principal-components analysis is one way to do this.

⁸The boldface numbers in Table I indicate each variable's highest factor loading. Factor identification is determined by examining only that subset of variables whose highest loading is on each factor.

Table 1. Principal Components Analysis (*n* = 6088)^a

Variable	Component 1	Component 2	Component 3	Component 4
Self-esteem	0.05	0.71	0.19	0.05
Values orientation	−0.20	−0.70	−0.26	−0.03
Attachment to school	0.04	0.25	0.76	0.08
Attachment to teachers	0.22	0.02	0.69	−0.01
School commitment	0.23	0.20	0.75	0.18
Perceived opportunities	0.00	0.57	0.47	0.09
Consequences of arrest	0.07	−0.49	0.07	0.07
School involvement	0.08	−0.04	0.10	0.83
Community involvement	0.10	0.04	0.07	0.83
Caring and trust	0.65	0.29	0.25	0.07
Identity support	0.41	0.62	0.09	0.01
Control and supervision	0.64	−0.06	0.26	−0.09
Intimate communication	0.79	0.02	−0.02	0.16
Instrumental communication	0.81	0.11	0.11	0.14

^aExtraction method: principal-components analysis. Rotation method: Varimax with Kaiser normalization.

subsequent cluster analysis.⁹

6.2. The Cluster Solution

Cluster analysis has been called the “black art” by statisticians, largely because of the many subjective decisions that must be made. One such decision is the identification of the optimal number of clusters. Some clustering algorithms are very obedient in the sense that they will provide whatever number of clusters one specifies. However, this particular number may not be optimal and the resulting solution can be entirely spurious. One of the critical early decisions therefore concerns the number of natural groupings (if any) that exist. Several guidelines are available to the analyst, including the percentage of variance explained (η), the dispersion among the groups, and the homogeneity within the groups (Everitt, 1993; Aldenderfer and Blashfield, 1984). Each is considered below.

The percentage variance explained (η) is the first statistic used to assess the “goodness of fit” of the clustering solution. The statistic measures the proportion of total variance that is attributable to the clustering solution.¹⁰ The statistic can be calculated for each cluster solution (two, three, four,

⁹One rule of thumb is that one should use component scores rather than any other form of data standardization as the basis for clustering if more than 60% of the total variance is explained by the significant principal components (Everitt, 1993).

¹⁰ η (percentage variance explained) is defined as $S^2_{\text{exp}} = 100 * (1 - SS_{\text{within}}/SS_{\text{total}})$.

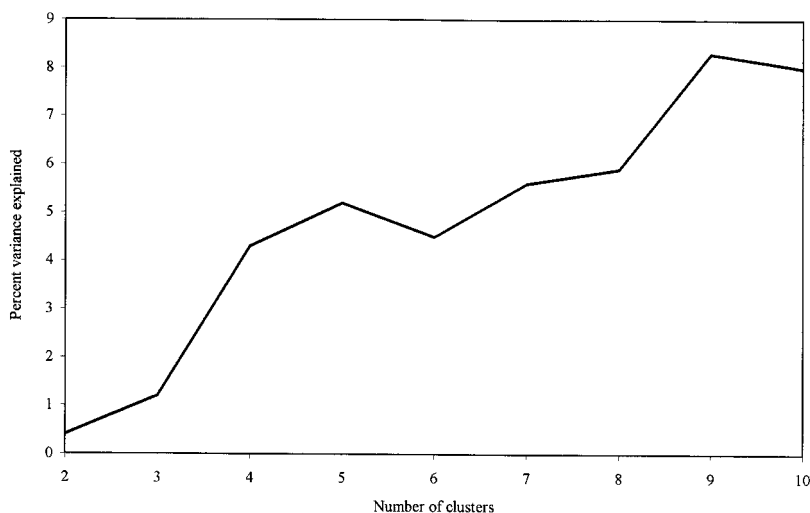


Fig. 1. Percentage variance explained for different cluster solutions.

etc., clusters) and the result can be graphed (see Fig. 1). In comparable fashion to evaluating the “scree slope” of a PCA, one assesses the competing cluster solutions by the identification of “knee points” in the distribution—the point at which there is a decline in the percentage variance explained from one cluster solution to the next.¹¹ As Fig. 1 indicates, the distribution of η was not ideal, and though no clear knee point emerged, the principal of parsimony clearly pointed to the five-cluster solution as optimal.¹²

The next stage in the analysis of the cluster solution involved consideration of cluster separation. A good-quality cluster solution should comprise clusters among which the separation (i.e., the distance between any pair of cluster centroids) exceeds 2 standard deviations from both group centroids. Table II presents the standard deviations for each cluster in the five-cluster solution.¹³ For each pair of clusters the distance between the centroids is

¹¹The absence of any “knee” may indicate that there are no natural groupings in the data.

¹²It is difficult to assess the quality of the overall solution with this statistic since it is very sensitive to total sample size and the total number of variables involved. Consequently, though an η of 10% may be good for one data set with a large number of variables and/or cases, it may be very poor for a data set with a small number of cases and/or variables.

¹³The data presented in this section represent the product of several additional analytic steps that cannot be fully described in a paper of this length. Briefly, the original five-cluster solution was examined for the presence of outliers (cases whose distance from the centroid exceeded two standard deviations). The outliers were temporarily removed and a new five-cluster solution was run. The centroid locations from this analysis were noted and then used as specified centers in a third five-cluster analysis utilizing all cases. The data presented are derived from this final five-cluster solution.

Table II. Distances Between Final Cluster Centers

Cluster	1	2	3	4	5	SD
1	—	2.25	2.34	2.50	2.20	0.57
2		—	2.16	2.28	1.88	0.47
3			—	2.39	2.06	0.55
4				—	2.18	0.58
5					—	0.41

significantly larger than the 2 standard deviation guide, suggesting that the five-cluster solution performs very well on the second diagnostic measure of separation.¹⁴

The third measure of the quality of a cluster solution is the homogeneity index (H_k).¹⁵ This measures the degree to which the variance *within* a particular cluster is smaller than the variance for the *entire population*. The index varies (usually) from 0 to 1, where a value of 1 indicates that the cluster is no more homogeneous than the data set before clustering, and 0 indicates that the cluster is perfectly homogeneous (i.e., all cases in the cluster have exactly the same value on each variable). As with η , the homogeneity index is sensitive to the number of variables and the number of cases in the cluster, making it difficult to compare group homogeneity from one clustering solution to the next. In practice, a homogeneity index below 0.25 is considered excellent, below 0.5 good, and below 0.75 acceptable. Using these guidelines clusters 5 and 2 were good and the remaining three clusters all had homogeneity index scores that were well inside the acceptable range.

Together the three diagnostic measures of variance explained, cluster separation, and cluster homogeneity all support the conclusion that the five-cluster solution derived is statistically sound. The next test is to determine its stability.

6.3. Stability of the Cluster Solution

A poor cluster solution will tend to shift its structure even if the same data are entered in a different order of cases. One test of cluster stability is therefore to reorder the cases, repeat the analysis, and compare the results. In addition to this technique, we utilized a more rigorous approach of drawing a series of bootstrap samples from the full data set. A set of four randomly selected subsamples of 25, 50, 75, and 99% of the total sample was

¹⁴For example, the distance from centroid 1 to centroid 2 is 2.25 when twice the standard deviation for centroid 1 ($2 * 0.57$) and for centroid 2 ($2 * 0.47$) combines to 2.08.

¹⁵The homogeneity index (H_k) = $(SS_{Gpk}/n_k)/(SS_{total}/N)$. This measure is not routinely available from SPSS output and must be calculated separately. The appropriate syntax for the calculation is available from the first author.

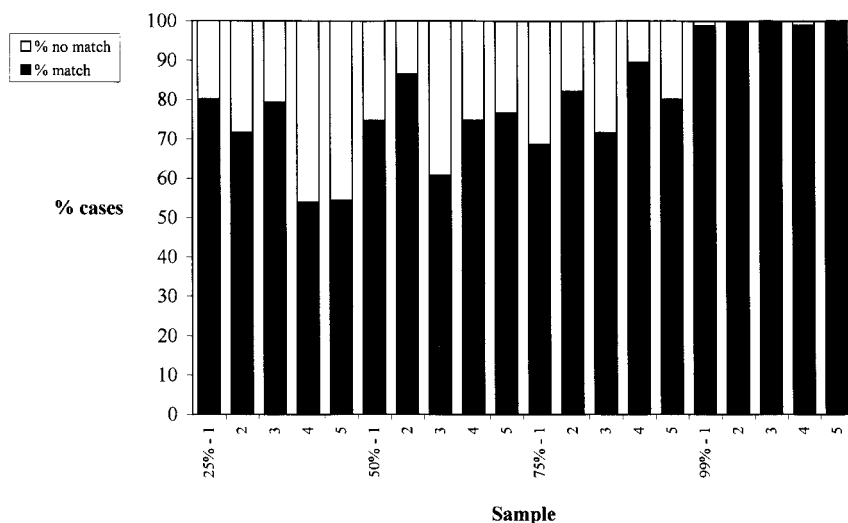


Fig. 2. Testing for cluster stability—comparing the bootstrap samples.

drawn, and for each, a five-cluster solution was independently derived. This process was repeated four times, to yield 20 subsamples. The cluster membership of cases within each of the 20 random subsamples was then compared to the membership of the original, total-sample, five-cluster solution. Though expectations of a perfect match are unrealistic, it is reasonable to assume that a cluster solution identifying “natural” clusters of juveniles should be replicated reasonably well within each random subsample. As with so many aspects of cluster analysis, there is no standard approach to this problem and no established criteria exist for assessing the replication of the cluster structure exists. Figure 2 presents the results for each of the bootstrap samples. Overall, the samples matched the original cluster structure in about 85% of cases, with the range varying from 53 to 99%. The results show that about 15% of cases change their cluster membership upon validation testing. In interpreting this figure it should be remembered that the cluster analysis classified *all cases*, irrespective of whether or not they are well suited to a cluster. Therefore all the clusters contain some unique “outlier” cases that do not fit well into any group and may well “float” from being an outlier in one cluster to being an outlier in a different cluster as the solution changes. Thus, in any series of cluster analyses we must expect some of the more peripheral members of the cluster groups to be reallocated as the boundaries of the clusters change. The key question is whether or not the core of each cluster (as defined by the centroid) shifts perceptibly from one solution to the other. Separate analyses (not presented

Table III. Distance of the Case from Its Classification
Cluster Center

Cluster	Mean	SD	Variance	<i>N</i>
1	1.47	0.59	0.35	1056
2	1.31	0.50	0.25	1156
3	1.52	0.54	0.29	844
4	1.43	0.56	0.31	1050
5	1.20	0.42	0.17	1708
Total	1.36	0.53	0.28	5814

here) show that such change did not occur during the bootstrap validation.

To this point we have identified five clusters and indicated that individual cases can be near or distant from the cluster centroid. Consequently, individual clusters can vary in terms of their internal structure with varying proportions of cases proximate to or distant from their respective cluster centroid. Table III presents one view of the cluster showing how the mean distances from individual case to cluster centroid vary. The results show that cluster 5 (the modal cluster) has the smallest average distance to centroid, followed by cluster 2, and the remaining three clusters are all more or less comparable.

6.4. Fine-Tuning the Solution

The initial five-category cluster solution included all cases from the sample. As noted above, a small number of the cases are unusual and do not fit well in any particular cluster. These outliers can be identified in cluster analysis by their large distance from the cluster centroid. In the present analysis these outliers¹⁶ were temporarily removed from the data set and a new five-cluster solution was derived. Omitting the outliers allows for the identification of cluster centroids whose location is not disproportionately affected by the large distances of a few outliers. Once the new cluster centroids are identified they become the “seeds” for a final cluster analysis that includes all cases.¹⁷

6.5. Describing the Clusters

Although the clusters were created on the basis of the four principal components defined in Table I, it is difficult to characterize each cluster on the basis of composite component scores. Consequently we use the mean

¹⁶The outliers were defined as cases whose distance exceeded the mean by 2 standard deviations.

¹⁷This final analysis does not allow the cluster solution to respecify the cluster centers, effectively “anchoring” the solution by the centroids identified without outliers.

cluster scores on each of the 14 original basis variables to describe the five clusters (Table IV).¹⁸ Based on these and other descriptor variables such as age, race, gender, family, and offense histories (measures not used to develop the clusters but likely related to the classification), we have identified the key characteristics of each of the five groups.

We are aware of the risk involved in labeling the types—the label may serve to mislead decision-makers about the treatment needs of an individual youth. However, the typology will undoubtedly be of more value if users can communicate information about groups of youths by some means that does not require extensive memorization. We have therefore provided labels to these types that reflect our current understanding of their characteristics.

6.5.1. Cluster 1—Autonomy Seeking

This group comprises older, mostly male youth who have not experienced negative family situations but reject attempts by parents to exert control. These youth are uninvolved with school and community resources of any kind and use drugs as a means of dealing with internal conflict. These youths do not have an optimistic view of their future and are not particularly open to treatment but, rather, wish to think that they have things under control. The key features of the group are

- very antisocial values,
- below-average self-esteem,
- very low school bonding scores (except for consequences of arrest), and
- below-average family bonding scores.

In addition, this cluster tends to have

- a disproportionately high percentage of Hispanic and White juveniles and a disproportionately low percentage of Black juveniles,
- the lowest family closeness measures and the highest proportion of juveniles reporting family problems at home, and
- the highest assessed overall need, including the highest need for education, peer relations, and sex education.

6.5.2. Cluster 2—Neurotic Anxious

This type of youth feels extremely rejected by his or her parents, is angry and hurt, experiences some degree of self-blame, and has had good experiences with teachers and school. The key features of the group are

¹⁸A more in-depth discussion of the types and the commonality between this typology and other existing typologies is given by Harris and Jones (1999).

Table IV. Cluster Scores on Basis Variables

Variable	Cluster					Total
	1	2	3	4	5	
Self-esteem						
Mean	7.73	8.63	6.47	8.73	9.29	8.34
SD	2.13	1.51	2.14	1.44	0.95	1.87
Values						
Mean	20.71	17.55	22.61	16.05	12.11	17.07
SD	6.73	5.80	5.33	5.97	5.33	6.90
School attachment						
Mean	2.60	4.25	3.87	4.14	4.38	3.93
SD	0.95	0.69	1.02	0.86	0.69	1.03
Attachment to teachers						
Mean	2.68	3.73	3.96	3.74	4.00	3.67
SD	0.86	0.74	0.72	0.83	0.67	0.88
School commitment						
Mean	2.54	3.44	3.40	3.61	3.70	3.40
SD	0.55	0.55	0.59	0.54	0.52	0.67
Perceived opportunities						
Mean	3.06	4.05	3.11	4.04	4.29	3.81
SD	1.01	0.76	1.05	0.85	0.69	0.99
Consequences of arrest						
Mean	3.04	2.98	3.84	3.08	2.69	3.07
SD	1.14	1.19	1.08	1.20	1.21	1.23
School involvement						
Mean	1.26	1.43	1.67	3.48	1.38	1.85
SD	0.62	0.82	0.98	1.22	0.71	1.23
Community involvement						
Mean	1.63	1.73	1.90	3.84	1.78	2.19
SD	1.00	1.01	1.14	1.03	1.02	1.34
Caring and trust						
Mean	3.55	3.46	3.94	4.11	4.36	3.93
SD	0.69	0.77	0.68	0.59	0.46	0.72
Identity support						
Mean	3.12	2.99	2.78	3.54	4.05	3.38
SD	0.84	0.90	0.84	0.85	0.70	0.95
Control and supervision						
Mean	3.25	3.05	3.93	3.62	4.03	3.60
SD	0.80	0.87	0.68	0.83	0.69	0.86
Intimacy and communication						
Mean	2.32	1.56	2.82	2.85	3.03	2.54
SD	0.96	0.62	1.03	1.03	0.90	1.07
Instrumental communication						
Mean	2.50	2.02	3.15	3.26	3.47	2.92
SD	0.90	0.71	0.98	0.89	0.76	1.01
Total <i>N</i>	928	1210	919	1168	1581	5806

- normal self-esteem and values scores,
- slightly above-average scores on most school bonding subscales but low scores on both involvement scales, and
- very low family bonding scores (the lowest average scores on four of the five subscales).

In addition, this cluster tends to have

- a high percentage of females,
- high levels of family violence,
- the highest proportion of first-time offenders,
- older juveniles—more juveniles who are above-average, and
- more juveniles with an above-average age at first arrest.

6.5.3. Cluster 3—Alienated Deprived

These juveniles are very antisocial and have very low self-esteem. Though they are the most likely to believe that their arrest is damaging to their future and have little hope for their future life-chances, they tend to score around the average on most of the other school and family bonding subscales. The one exception is identity support, indicating that they do not believe that their family is supportive of the type of person they have become. The key features of the group are

- very low self-esteem and very antisocial values,
- school bonding scores that are generally around average except for “perceived opportunities” and “consequence of arrest,” and
- family bonding scores that are generally above average (only “identity support” shows a low family bonding score).

In addition, this cluster tends to have

- a disproportionately large percentage of Hispanic juveniles and low percentage of White juveniles,
- high levels of family violence,
- above-average levels of weapons use in instant offense,
- younger juveniles—more juveniles who are below-average age, and
- more juveniles with a below-average age at first arrest.

6.5.4. Cluster 4—Passive Conformist

This group of juveniles is characterized by a lower average age, implying immaturity, high scores on scales pertaining to positive relationships with others, and the highest scores on school and community involvement. The key features of the group are

- average scores for self-esteem and values,

- school bonding scores that are generally above average—though the two involvement measures (school and community) are very high for this group, and
- family bonding scores generally at or above average.

In addition, this cluster tends to have

- a disproportionately low percentage of Hispanic juveniles and high percentage of Black juveniles,
- the highest ratings of family closeness,
- the lowest proportion of females, and
- the highest proportion of juveniles with prior arrests and recent prior arrests.

6.5.5. Cluster 5—*Detached Defended*

This group of youths is the most positive in terms of almost all school and family bonding scales. In addition, they have the highest self-esteem and the most prosocial values. As with most other groups they have low scores for school and community involvement. Despite the positive image there are certain other characteristics of this group that suggest caution. They are one of the most likely to have delinquency histories and to have siblings that have been arrested. They are the most likely to have injured a victim in the commission of their offense.

As well as being the largest single group for this population of delinquent youths, cluster 5 is the most common type of youth in programs for violent youth with mental health issues. Combining the unusually high mean score on “perceived opportunities” for the future and the atypically low mean score on “consequences of arrest” with some of the other group characteristics suggests that these youths may be detached from reality and somewhat unaffected by their own behavior or the reactions of others.

The key features of the group are

- very high self-esteem and very prosocial values,
- generally very high school bonding scores—except for below-average scores on consequences of arrest, school, and community involvement, and
- uniformly very high or high family bonding scores.

In addition, this cluster tends to have

- the lowest levels of family violence and family problems,
- the highest level of family closeness,

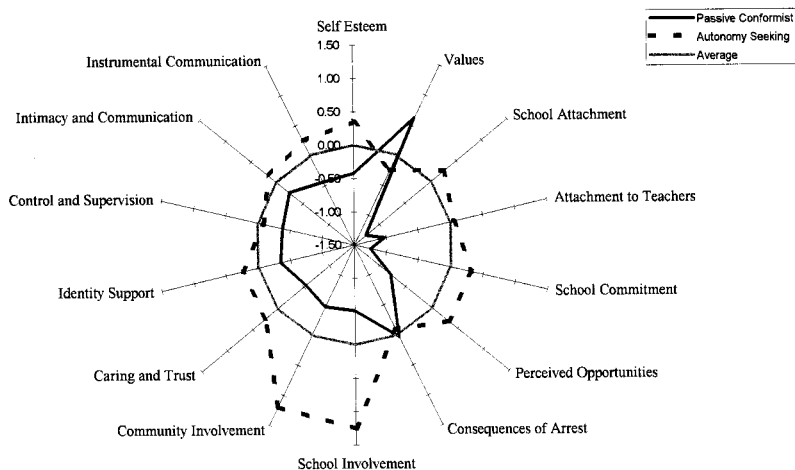


Fig. 3. Visualizing multidimensional clusters—a comparison of “autonomy seeking” and “passive conformist” juveniles.

- the least likelihood to have juveniles with alcohol/drug abuse problems or to have mothers and/or fathers with alcohol and/or drug abuse problems,
- the lowest proportion assessed as high need,
- the lowest proportion assessed as high risk,
- older juveniles—more juveniles who are above-average age, and
- more juveniles with an above-average age at first arrest.

Figure 3 shows how one can visually compare the clusters simultaneously on all basis variables. The two types selected—clusters 1 (autonomy seeking) and 4 (passive conformist)—represent the most diverse pairing since they have the largest overall intercentroid distance.¹⁹ Figure 3 shows the standardized scores for the two types and illustrates the extent to which cluster 1 scores far higher than cluster 4 on almost all measures (the one exception is the values dimension, where “high” scores indicate more anti-social values). The only area of similarity for the two types is the “consequence of arrest” subscale, where both clusters scored around the average for the population as a whole.

6.6. Validity of the Cluster Solution

Our focus to this point has been an assessment of the quality and stability of the cluster solution as well as an understanding of the character of

¹⁹The radar chart shows the z-scores for all 14 original subscales.

each cluster. However, it is equally important that we validate the typology before using it as part of the larger evaluation structure.

To assess the validity of the typology we adopted two approaches. In the first we compared the current empirically derived typology with several existing classification systems that have been and/or are still being used in juvenile delinquency research. A detailed comparison of the typologies is presented elsewhere (see Harris and Jones, 1999). Here we examine the extent to which the *ProDES* typology includes the key theoretical constructs embedded in the leading juvenile classification systems—the I-Level (Warren, 1976; Harris, 1988), Moral Development (Jennings *et al.*, 1983), Conceptual Level (Brill, 1978), Quay (1979), and MMPI-based (Zager, 1979; Megargee, 1977; Jesness, 1983) systems. Van Voorhis (1994) identified the following eight common constructs among the five systems.

- *Cognitive Development*: I-Level, Moral Development, and Conceptual Level each specify continua of psychological development. While the relatively low age of *ProDES* type 3 suggests lower maturity than the other types, the *ProDES* typology contains no developmental dimension.
- *Neurotic*: The I-Level (here including Warren and Jesness), Quay, and MMPI-based systems identify persons whose behavior is attributed to substantial internal conflict. Typically, these persons are introspective and anxious. Within the *ProDES* typology, internal conflict can be inferred from a combination of low self-esteem and low scores on “caring and trust” and “identity support” (from the family bonding scale). Juveniles of types 1 and 2 are the most likely to display these characteristics.
- *Situational*: The I-Level, Quay, and MMPI-based systems describe some individuals as being involved in the justice system as a result of reactions to immediate situations in their lives. Van Voorhis (1994) reports finding no relationship among the classifications of the three typologies on this type in an adult population, suggesting that they are often used as a catch-all for persons who cannot be classified. *ProDES* is unable to tap this dimension.
- *Immature Dependent*: The I-Level, Quay, and MMPI-based systems contain types who are identified as passive, compliant, fearful, and dependent. This construct is not measured directly by *ProDES*. We have found, however, that certain patterns of scores indicate a need to please others and that this pattern is consistent with staff descriptions of youths classified as type 4.
- *Character Disorder, Psychopathy, Sociopathy*: The I-Level, Quay, and MMPI-based systems contain types that are manipulative,

charming, and impulsive. I-Level makes no specific reference to psychopathy or sociopathy. Moreover, *ProDES* contains no measures of this dimension. Based on our exploration of the data we believe that an atypically high score on “perceived opportunities” combined with an atypically low score on “consequences of arrest” may indicate sociopathy, especially if combined with an extensive history of serious delinquency, high self-esteem, and moderately prosocial values.

- *Subcultural*: While the MMPI-based system contains no such type, both I-Level and Quay identify persons who have been socialized into a delinquent subculture and who seek to adhere to the standards of that group. This dimension is measured by means of the values orientation scale.
- *Power-Oriented*: Persons who seek to counteract the authority of others and gain control of situations themselves are found in all three typologies. This dimension is measured by the values orientation scale.
- *Manipulative*: All three systems identify persons who seek to manipulate others for no particular reason. This is their primary means of social interaction. *ProDES* does not measure this dimension. Weak attachments to others, however, gives us some indication that manipulation is a likely response mode, and attachments to others are measured by the school and family bonding scales.

Keeping these constructs in mind as we compare the *ProDES* types to the other typologies ensures that we do not drift far from the concepts that underlie the other typologies. However, it is equally important to recognize where the *ProDES* typology uses constructs that differ from those of the other typologies. Elsewhere (Harris and Jones, 1999) we have specified hypotheses about each type, including its similarity to those of other typologies and its criminogenic or responsivity needs. We have also directly compared each of the *ProDES* types to the categories developed in other classifications.

The second approach to assessing the validity of the typology was to “ground” the results in the experiences of those persons working in the field, those who know the juveniles first-hand. To test whether or not the cluster solution made intuitive as well as statistical sense, we invited three staff members from a respected delinquency treatment program to review the names of juveniles from their own program that our analysis had placed together in each of the clusters. The staff members were told only that there were five groups and that the classification was based upon the self-reported information collected by *ProDES*. They were asked to describe the main

characteristics of each group based upon their knowledge of the juveniles placed there by the cluster analysis. They were also asked to identify any juveniles (if any) who did not belong in the group. These staff were told that they could discuss the juveniles in terms of not only the basis variables (the self-report measures) but also any other key characteristics that they felt discriminated among the groups (e.g., drug and or alcohol use, violence, educational ability etc.).

Working only from the names in each group,²⁰ the staff discussed their images of each juvenile. The discussion lasted several hours, and for each of the five groups the staff independently identified the majority of the characteristics empirically derived from the statistical analysis—in terms of both the basis and some descriptor variables. For example, our notes show that they identified juveniles in cluster 3 as having low self-esteem and anti-social values and those in cluster 2 as having particularly poor family relationships.

The fact that the program staff, with their personal, often intimate knowledge of each of the youths, were so closely aligned with our independent and empirically based classification convinced us not only that the classification was statistically adequate but that it was grounded in the experience of program life.

7. UTILITY

Megargee (1977) discussed several goals for a good classification system. Among the more important criteria were

- (1) sufficient completeness so that most offenders can be classified,
- (2) clear operational definitions to avoid ambiguity,
- (3) adequate reliability across raters,
- (4) sufficient validity with respect to the attributes and behaviors predicted by the systems,
- (5) sufficient dynamic properties so that a change in attributes, behavior, or status is reflected by a change in classification,
- (6) implications for treatment or intervention, and
- (7) economical classification of large numbers of subjects.

We regard these criteria as sound for judging the work we report here. The typology we have developed has been shown to be reliable, valid, and stable (at least in the short term). Further, it is exhaustive and efficient to use. The typology also has several characteristics often absent from the more established systems. The typology is dynamic in the sense that it can adapt

²⁰Each cluster had between 15 and 25 names.

to changes in the population being classified. In Philadelphia we are finding that juveniles entering the system in 1997 and 1998 are far more likely to have histories of mental health problems, to have drug and alcohol abuse problems, and to have parents with substance abuse problems. As the nature of the juvenile population changes, it seems increasingly inappropriate to continue to classify them using a system developed and normed on a relatively small, dated, and substantively different sample.

Because of the way *ProDES* is constructed—collecting data at program intake and discharge—it is possible (indeed likely) that cluster membership will change over time even for a particular individual. In most cases program interventions specifically seek to bring about personality changes as a means of reducing the propensity for future delinquency. If program intervention is to be successful, we would expect to find changes on multiple personality measures—higher self-esteem, less antisocial values, and more positive school and family bonding scores. If the juvenile changes substantially on several of these measures, it is possible that their cluster membership will also change. One of the strengths of the present classification approach is its ability to monitor short-term change on an individual level (Do individuals change cluster membership from intake to discharge?) as well as longer-term shifts in the juvenile population as a whole.

The ability to monitor individual change over time is important since it fits theoretically with the development of outcome-based information systems in which dynamic measures seem superior to the more traditional static measures (see Bonta, 1996). We agree with the general consensus that measures of change during a program intervention are more likely to enhance our understanding of the treatment than are static measures either at the point of intake or at discharge.

The fact that a classification can change over time may not, at first sight, seem to be a good thing. Certainly, a typology that changes significantly in short periods of time would be of little practical value either for treatment or prediction. However, the juvenile population is constantly changing, and it seems equally illogical to believe that a typology developed decades ago on the basis of several hundred juveniles will have universal application over extended periods of time. The most realistic perspective is one in which the typology is consistently being tested and, if necessary, modified every few years.²¹ Information systems such as *ProDES* make such an approach possible. Despite the difficulties this will cause for the interpretation of juvenile personality types (the boundaries of the “types” may keep shifting), it would seem preferable to the continued use of obsolete classification systems that no longer adequately describe the juvenile population.

²¹One possibility would be to adopt a smoothing approach in which change is ameliorated across successive iterations.

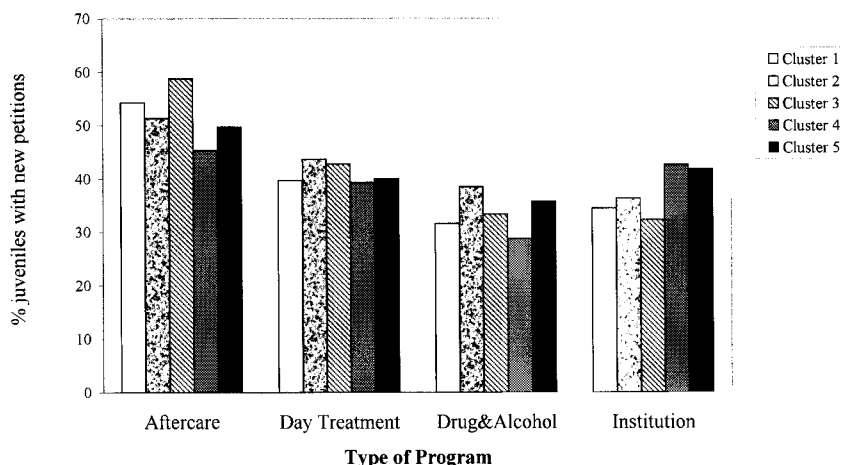


Fig. 4. Visualizing multidimensional clusters—a comparison of “autonomy seeking” and “passive conformist” juveniles.

The present study is not the first to take this approach to the question of classification. Simourd *et al.* utilized a similar approach with a much smaller sample of cases focusing on risk/needs psychometric measures. In both studies the combination of theoretically sound measures with appropriate statistical analyses produced five-cluster typologies that stood up to rigorous tests of reliability and validity. Unlike that of Simourd *et al.*, the present typology was not developed specifically as a risk instrument, though the variables upon which the typology was developed were all identified as important psychological indicators of delinquency.

We tested the present typology against population reoffending data and found little discrimination among the cluster groups. It appears that the different types of juveniles do not vary at all in their overall propensity to reoffend. However, Palmer (1994) has argued that interventions are optimized when we appropriately match the needs of the juvenile with the services of the program. This suggests that there may be an interaction between the type of juvenile and the type of program. For example, the same type of juvenile may have different outcomes when they enter different types of program. We test this assumption in Fig. 4, which shows that the proportion of juveniles who reoffend does vary by juvenile type within program type.²² For example, type 3 juveniles were the most likely to reoffend during and

²²The reoffending figures comprise the time spent in the program plus the first 6 months following discharge. Reoffending is measured in terms of new petitions before Family Court rather than arrests. Despite the obvious shortcomings of the court measure, we found that in Philadelphia they are vastly more reliable than police statistics.

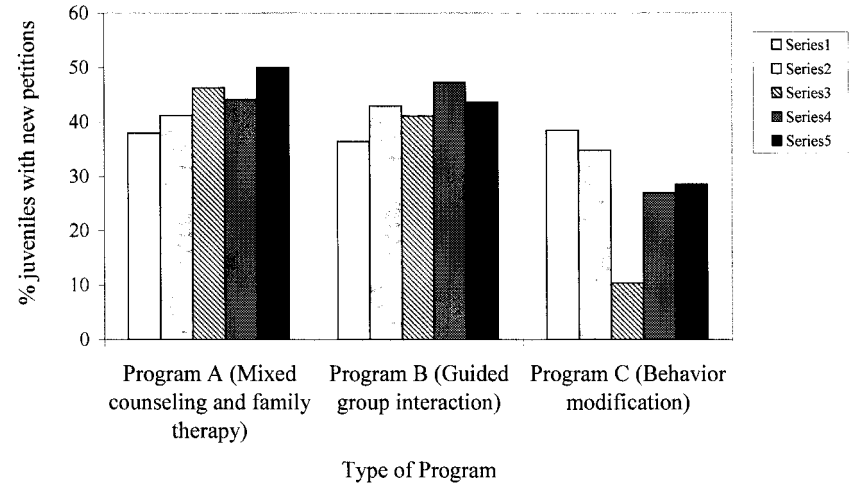


Fig. 5. Juvenile reoffending by cluster and type of program.

after aftercare programs but the least likely to offend during and after placement in an institutional program. In contrast, cluster 4 has the highest rate of new petitions among juveniles in institutional placements but the lowest rate in each of the other three program types.

Figure 4 also shows that the type of juvenile is of less importance to reoffending rates when we are dealing with day treatment than with after-care providers. Thus, the interaction effect between type of juvenile and type of program is not the same for all program types—in some instances it seems to matter a great deal; in others it is of less importance.

From our experience with the Philadelphia system we know that even within a particular program type, there are distinct variations in the treatment modality. We therefore selected three specific institutional programs that are characterized by different primary treatment modalities—Program A represents a mixed counseling and family therapy approach, Program B a guided group interaction model, and Program C a behavior modification approach. Figure 5 shows that at the level of the individual program there is strong evidence for the interaction effect—type 1 juveniles seem to have comparable reoffending figures irrespective of the particular treatment modality of the program. In contrast, type 3 juveniles do a lot better in the behavior modification program than in either of the other two modalities.

8. CONCLUSION

Current research in risk/needs assessment and prediction emphasizes the use of dynamic or change-oriented measures. We believe that the future

of juvenile classification systems requires a similar ability to develop dynamic models that are capable of changing as the nature of the population itself changes. Existing classification systems are generally developed on the basis of relatively limited, time-bound samples. Such systems tend to have a "life cycle" in which they are developed, used widely, and then tend to fall into disuse largely because they are seen as increasingly obsolete and out of touch with the reality of contemporary juvenile personality.

We have developed an approach that we feel can be usefully replicated. The selection of the basis variables is grounded in both the delinquency literature and the operational practices of juvenile delinquent programs. The data are collected as part of an ongoing evaluation-based information system rather than a single or series of discrete cross-sectional studies. The statistical analysis combines data reduction (through PCA) and clustering techniques to derive empirically both the number of groups and their internal structure. The classification system should be periodically reevaluated and its validity and reliability reestablished using, where possible, a triangulated approach of quantitative and qualitative techniques.

The result is a dynamic approach to classification in which the typology itself can adjust to shifts in the juvenile population's scores on the basis variables. If there should be some marked change in the distribution of scores on one or more of these measures, it may affect the latent factor structure of the PCA, the number of discrete groups, and even the definition of the groups themselves. In other words, the classification may change if the underlying composition of the juvenile population itself changes.

Of course, the development of the juvenile typology is only the beginning of the evaluative process. By controlling for the type of juvenile our ultimate goal is to develop a more finely tuned understanding of the age-old questions surrounding differential treatment. The research reported in this paper shows that we are moving close to the time when we can better answer the question rhetorically posed by researchers such as Andrews (1990), Palmer (1992), and Van Voorhis (1997)—What works, with whom, and in what circumstances?

REFERENCES

- Aldenderfer, M. S., and Blashfield, R. K. (1984). *Cluster Analysis*, Quantitative Applications in the Social Sciences, No. 44, Sage, Thousand Oaks, CA.
- Andrews, D. A., and Bonta, J. (1994). *The Psychology of Criminal Conduct*, Anderson, Cincinnati, OH.
- Andrews, D. A., Zinger, I., Hoge, R. D., Bonta, J., Gendreau, P., and Cullen, F. T. (1990). Does correctional treatment work? A clinically relevant and psychologically informed meta-analysis. *Criminology* 28(3): 369–404.

- Baird, S. C. (1986). Classifying juveniles—An important management tool. In Dziobek, L. *et al.* (eds.), *Issues in Juvenile Delinquency*, American Correctional Association, Laurel, MD.
- Bonta, J. (1996). Risk-needs assessment and treatment. In Harland, A. (ed.), *Choosing Correctional Options That Work: Defining the Demand and Evaluating the Supply*, Sage, Thousand Oaks, CA.
- Brannon, J., *et al.* (1989). Applied behavioral typologies: Implications for the residential treatment and community re-integration of juvenile offenders. *J. Offender Counsel. Serv. Rehab.* 14(3): 77–87.
- Brennan, T. (1987). Classification: An overview of selected methodological issues. In Gottfredson, D. M., and Tonry, M. (eds.), *Prediction and Classification: Criminal Justice Decision Making*, University of Chicago Press, Chicago, pp. 201–248.
- Brill, R. (1978). Implications of the conceptual level matching model for the treatment of delinquents. *J. Res. Crime Delinq.* 15(3): 229–246.
- Cernkovich, S. A., and Giordano, P. C. (1992). School bonding, race and delinquency. *Criminology* 30: 261–291.
- Clements, C. B. (1996). Offender classification: Two decades of progress. *Crim. Just. Behav.* 23(1): 121–143.
- Cloward, R. A., and Ohlin, L. E. (1960). *Delinquency and Opportunity: A Theory of Delinquent Gangs*, Free Press, New York.
- Dembo, R. *et al.* (1984). Development and assessment of a classification of high risk youths. *J. Drug Issues* 24(1, 2): 25–54.
- Everitt, B. S. (1993). *Cluster Analysis*, Edward Arnold, London.
- Giordano, P. C., and Cernkovich, S. A. (1987). Family relationships and delinquency. *Criminology* 25: 295–321.
- Glaser, D. (1987). Classification for risk. In Gottfredson, D. M., and Tonry, M. (eds.), *Prediction and Classification: Criminal Justice Decision Making*, University of Chicago Press, Chicago, pp. 249–292.
- Goring, C. (1913). *The English Convict*, HMSO, London.
- Gottfredson, D. M. (1987). Prediction and classification in criminal justice decision making. In Gottfredson, D. M., and Tonry, M. (eds.), *Prediction and Classification: Criminal Justice Decision Making*, University of Chicago Press, Chicago, pp. 1–20.
- Gottfredson, S. D. (1987). Prediction: An overview of selected methodological issues. In Gottfredson, D. M., and Tonry, M. (eds.), *Prediction and Classification: Criminal Justice Decision Making*, University of Chicago Press, Chicago, pp. 21–52.
- Harris, P. W. (1988). The Interpersonal Maturity Classification System: I-level. *Crim. Just. Behav.* 15(1): 59–77.
- Harris, P. W., and Jones, P. R. (1999). Differentiating delinquent youths for program planning and evaluation. *Crim. Just. Behav.* (in press).
- Hirschi, T. (1969). *Causes of Delinquency*, University of California Press, Berkeley.
- Huizinga, D., Esbensen, F., and Weiher, A. W. (1991). Are there multiple paths to delinquency? *J. Crim. Law Criminol.* 82(1): 83–118.
- Jefferson, T. W., and Johnson, J. H. (1991). Relationship of hyperactivity and sensation seeking to delinquent subtypes. *Crim. Just. Behav.* 18(2): 195–201.
- Jennings, W., Kilkenney, W., and Kohlberg, L. (1983). Moral development theory and practice. In Laufer, W., and Day, J. (eds.), *Personality Theory, Moral Development and Criminal Behavior*, D. C. Heath, Lexington, pp. 281–355.
- Jesness, C. F. (1983). *The Jesness Inventory: Manual*, Multi-Health Systems, North Tona-wanda, NY.
- Jones, P. R. (1996). Risk prediction in criminal justice. In Harland, A. (ed.), *Choosing Correctional Options that Work: Defining the Demand and Evaluating the Supply*, Sage, Thousand Oaks, CA.

- Jones, P. R., and Harris, P. W. (1997). *ProDES, an Assessment of the Self Reported Scales*. Monograph, Crime and Justice Research Institute, Philadelphia.
- MacKenzie, D. L., Posey, C. D., and Rapaport, K. R. (1988). A theoretical revolution in corrections: Varied purposes for classification. *Crim. Just. Behav.* 15(1): 125–136.
- Megargee, E. I., and Bohn, M. J., Jr., with Myer, J., and Sink, F. (1979). *Classifying Criminal Offenders*, Sage, Beverly Hills, CA.
- Mezzich, A. C., Coffman, G., and Mezzich, J. E. (1991). Typology of violent delinquent adolescents. *J. Psychiatry Law* 92(2): 63–78.
- Miller, W. (1958). Lower-class culture as a generating milieu of gang delinquency. *J. Soc. Issues* 42(1): 5–19.
- Palmer, T. (1984). Treatment and the role of classification: A review of basics. *Crime Delinq.* 30(2): 245–267.
- Palmer, T. (1992). *The Re-Emergence of Correctional Intervention*, Sage, Thousand Oaks, CA.
- Quay, H. C. (1979). Classification. In Quay, H. C., and Werry, J. S. (eds.), *Psychological Disorders of Childhood*, 2nd ed., John Wiley and Sons, New York, pp. 1–42.
- Robinson, J. P., and Shaver, P. R. (1973). *Measures of Social Psychological Attitude*, Institute for Social Research, University of Michigan, Ann Arbor.
- Rosenberg, M. (1965). *Society and the Adolescent Self-Image*, Princeton University Press, Princeton, NJ.
- Sechrest, L. (1987). Classification for treatment. In Gottfredson, D. M., and Tonry, M. (eds.), *Prediction and Classification: Criminal Justice Decision Making*, University of Chicago Press, Chicago, pp. 293–322.
- Simourd, D. J., Hoge, R. D., Andrews, D. A., and Leschied, A. W. (1994). An empirically-based typology of young offenders. *Can. J. Criminol.* 36(4): 447–454.
- Smith, M. F. (1989). *Evaluability Assessment: A Practical Approach*, Kluwer Academic, Norwell, MA.
- Van Voorhis, P. (1994). *Psychological Classification of the Adult Male Prison Inmate*, State University of New York Press, Albany.
- Van Voorhis, P. (1997). An overview of offender classification systems. In Van Voorhis, P., Braswell, M., and Lester, D. (eds.), *Correctional Counseling and Rehabilitation*, 3rd ed., Anderson, Cincinnati, OH.
- Warren, M. Q. (1965). Implications of a typology of delinquents for measures of behavior change: A plea for complexity. *Youth Auth. Q.* 18(3): 1–7.
- Warren, M. Q. (1971). Classification of offenders as an aid to efficient management and effective treatment. *J. Crim. Law Criminol. Police Sci.* 62(2): 239–258.
- Warren, M. Q. (1976). Applications of interpersonal theory to offender populations. In Rosenheim, M. (ed.), *Pursuing Justice for the Child*, University of Chicago Press, Chicago.
- Warren, M. Q., and Hindelang, M. J. (1979). Differential explanation of offender behavior. In Toch, H. (ed.), *Psychology of Crime and Criminal Justice*, Waveland Press, Prospect Heights, IL.
- Wholey, J. S. (1994). Assessing the feasibility and likely usefulness of evaluation. In Wholey, J. S., Hatry, H. P., and Newcomer, K. E. (eds.), *Handbook of Practical Program Evaluation*, Jossey-Bass, San Francisco.
- Wilkins, L. T., and MacNaughton-Smith, P. (1964). New prediction and classification methods in criminology. *J. Res. Crime Delinq.* 1: 19–32.
- Zager, L. D. (1979). The MMPI-based criminal classification system. *Crim. Just. Behav.* 15(1): 39–57.