J Quant Criminol (2008) 24:179–203
DOI 10.1007/s10940-008-9045-7

ORIGINAL PAPER

Towards an Explanatory Taxonomy of Adolescent Delinquents: Identifying Several Social-Psychological Profiles

Tim Brennan · Markus Breitenbach · William Dieterich

Published online: 29 February 2008

Abstract Taxonomic structure is examined in two large samples of delinquent youth in a domain of socio-psychological and personality factors. This paper offers a partial empirical test of the overlapping theoretical taxonomies of Moffitt (Pshycol Rev 100:674–701, 1993), Lykken (The antisocial personalities, 1995) and Mealey (1995). The first sample consisted of juvenile offenders ($n = 1{,}572$) from three state systems. Multiple cluster analysis methods were applied (Wards method, standard K-means, bootstrapped K-means and a semi-supervised pattern recognition technique). Core or exemplar cases were identified by means of a voting procedure. Seven clusters recurrently emerged across replications. While clear analogues of Moffitt's two main categories were found, several additional stable subtypes emerged that were clearly reminiscent of Lykken's sociopathic, neurotic-internalizing and "normal" types. However, boundaries between types were fuzzy and unstable, and many unclassified cases existed. Internal validation was assessed by cross-method verification. External validation assessed type differentiation on several delinquent behaviors. Finally, generalizability was assessed by repeating the clustering on a large replication sample ($n = 1{,}453$) from another state. Six of the seven initial types re-emerged.

Keywords Taxonomy · Patterns · Cluster analysis · Delinquent types · Bootstrap aggregation

T. Brennan (✉) · W. Dieterich
Northpointe Institute for Public Management Inc., P.O. Box 309, Traverse City, MI 49685, USA
e-mail: tbrennan@npipm.com

T. Brennan
Institute for Cognitive Science, University of Colorado, Boulder, CO, USA

M. Breitenbach
Department of Computer Science, University of Colorado, Boulder, CO, USA

Introduction and Background

The debate over the existence of criminal types remains controversial. A dominant position is that offender types based on differential etiological causal processes do not exist; instead, many prominent criminologists emphasize general theories of antisocial behavior (Gottfredson and Hirschi 1990; Jessor et al. 1991; Osgood et al. 1988). An alternative position is that types exist and can be both theoretically and empirically justified.

In this discussion it is important to distinguish between several approaches in recent taxonomic research in criminology. One strand of this work uses group-based methods to identify crime sequence trajectories that may reflect distinct criminal careers (Nagin and Land 1993). A second approach seeks to build explanatory taxonomies of offenders based on psychosocial, biological, personality and other explanatory factors (Harris and Jones 1999; Lykken 1995; Mealey 1995; Moffitt 1993; Van Voorhis 1988).The present study follows the second of these two approaches. It examines latent taxonomic structure in two large heterogenous samples of delinquent youth assessed on a wide range of psychosocial and risk factors previously linked to delinquent behavior.

Theory and Taxonomy in Delinquency

The taxonomic position suggests that the complexity of antisocial behavior cannot be reduced to a single general or universal theoretical process. It assumes that distinct etiologies exist and that the taxonomic approach may unravel these into understandable and coherent categories to more clearly reflect these diverging etiologies (Gibbons 1975; Huizinga et al. 1991; Jones and Harris 1999; Lykken 1995; Moffitt 1993; Paternoster and Brame 1997; Van Voorhis 1994; Zhang et al. 2002). In an influential publication, Lykken (1995) invoked the concept of "structural heterogeneity" to support qualitatively different offender types. His descriptions of the different causal processes underlying psychopathic and sociopathic types broadly overlap with Moffitt (1993) and Mealey (1995). In advocating the taxonomic approach to the explanation of crime Lykken (1995) commented as follows:

> There seems to be an almost irresistible tendency for criminological theorists to oversimplify the causes of crime, to underestimate the variety of psychological peculiarities that can contribute to the underlying dispositions for criminal behavior (p. 17).

With any theoretical taxonomy a key research challenge is to demonstrate the empirical reality, reliability and coherence of the proposed types. Moffitt (2003) refers to her main types as "theoretical prototypes." Lykken (1995) similarly acknowledges that his is an "armchair taxonomy" that clearly requires empirical verification. In both cases evidence for the validity and existence of these explanatory types remains inconclusive. Moffitt (2003) has enumerated several research gaps in her taxonomy, particularly the need for empirical elaboration and to establish the personal characteristics of her types. She also raised the issue of the hierarchical structure of her taxonomy by asking the question "Are two groups enough?" (p. 56). A finding of additional latent types would suggest that Moffitt's taxonomy may require further elaboration in order to more fully identify the heterogeneity among delinquent youth.

Another issue with strong theoretical implications is the controversy over categorical versus dimensional approaches to the representation and analysis of delinquency. This questions whether the underlying reality of offender causation is one of discrete groups or

continuous dimensional differences (Osgood 2005). This debate has also emerged in recent proposals for the evolution of the American Psychiatric Association's Diagnostic and Statistical Manual (American Psychiatric Association 1994; Beutler and Malik 2002). The dimensional approach is more clearly associated with the search for general theories of delinquency and has been the preferred approach of most criminological theorists. Osgood (2005) explicitly states his preference for general theories and "dimensional" conceptions of the underlying causal reality of delinquency (see also Sampson and Laub 2005). Farrington (2003) underlined this dominance by noting that most recent theoretical developments, with the exception of Moffitt's (1993) taxonomic theory, do not support the idea of types. Hirschi and Gottfredson (1994) starkly dismiss taxonomic heterogeneity—at both the behavioral and explanatory levels - and ascribe most criminal behavior to a single "persistent underlying trait" (see also Britt 1994).

However, a reliable demonstration of distinct types with diverging explanations of crime would challenge the proponents of a general theory of crime. Ultimately this debate may only be settled by careful demonstration of whether distinct and homogeneous latent clusters can be found and whether these exhibit explanatory coherence (Thagard 1992). We agree with Osgood (2005) that this matter is not yet settled for criminology and will ultimately be clarified by the empirical structures of our data.

Reasons for the Difficulty of Taxonomic Research in Criminology

The unfinished and controversial nature of typological research in delinquency is due to several factors. First, delinquency is characterized by extreme multivariate complexity across several interacting explanatory domains (Walsh 2002). Second, taxonomic structure and latent classes in delinquency may only be characterized by fuzzy boundaries and graded prototype structures, rather than well separated distinct types. It may not offer the strong taxonomic organization that characterizes the world of chemistry, physics or other harder sciences, where the "holy grail" of taxonomic research is exemplified in classifications such as the periodic table, with crisp boundaries, reliable identification criteria and coherent internal theories for each class (Bryant 2000). Third, taxonomic studies in delinquency have been plagued by methodological deficiencies. Many prior taxonomic studies have used inappropriate methods to detect types (for example, confirmatory factor analysis, structural equation modeling); small or inappropriate samples; over-simplified measurement domains; and inadequate validation designs. Lenzenweger (2004) offers a review of these and other misuses of taxonomic methods. Such misuses are unfortunate given advances in pattern recognition and taxometric methods, and in new validation designs (Milligan 1996; Wishart 2003).

Recent Progress in Taxonomic Research in Delinquency

In spite of these difficulties, clear progress has occurred regarding delinquent taxonomies in the last decade. First, three important theoretical taxonomies—from different disciplines—offer fairly similar descriptions (see below) of several categories of antisocial offenders (Lykken 1995; Mealey 1995; Moffitt 1993).

Second, a body of empirical results using a variety of taxometric methods is building support for the existence of delinquent typologies(Aalsma and Lapsley 2001; Harris and Jones 1999; Huizinga et al. 1991; Jefferson and Johnson 1991; Jones and Harris 1999; Mezzich et al. 1991; Nagin and Paternoster 2000; Potter and Jenson 2003; Skilling et al. 2001; Sorensen and Johnson 1996). Additionally, recent work in child developmental

psychology and developmental pathways also offers support for distinct trajectories leading to delinquency. Frick (2004) in summarizing this discipline concluded that there was evidence for distinct developmental pathways leading to conduct disorder and that these pathways involve unique causal processes (see also Loeber 1996; Loeber et al. 1997).

However, the above taxonomic studies rarely provide operational procedures for type identification, which has severely limited replication studies. Consequently, most have had limited empirical replication and rarely with appropriate offender samples, appropriate mathematical taxonomic methods or effective verification designs. Most of these recent papers consistently call for replication on larger and more clearly delinquent samples.

Goals

This paper has three primary goals in proposing a scientifically based taxonomy of juvenile offenders.

The first goal is to illustrate recent developments in taxonomic methods. We examine multiple verification across several taxonomic methods: classic K-means, bootstrapped aggregation K-means and a recently developed semi-supervised clustering method. The latter aims to discover intrinsic "natural" clusters in noisy or weakly structured data (Zhou et al. 2004). We also illustrate the cross-validation procedure of McIntyre and Blashfield (1980) to examine the generalizability of taxonomic findings.

Our second broad goal is substantive. We examine two large samples to assess the re-emergence of several delinquent types previously proposed in both classic and recent studies (Harris and Jones 1999; Lykken 1995; Mealey 1995; Megargee and Bohn 1979; Moffitt 1993; Van Voorhis 1994; Warren 1971). The basic profiles of these prior types are given below.

Third, we address the question of the "number of types" and the potential of a hierarchical taxonomy. Moffitt (2003) identified this question as an important issue for further research.

The basic structures of the most recurrent types in prior taxonomic literature are as follows:

1. *Normal or Situational Offenders*: This type is generally described as psychologically normal with few interpersonal relationship problems or risk factors and only minor delinquency. Delinquency may stem from accidental or stressful situational factors that overwhelm normal coping strategies (see Aalsma and Lapsley 2001; Huizinga et al. 1991; Lykken 1995; Van Voorhis 1994).
2. *Socialized Delinquents, Common Sociopaths or Subcultural Offenders*: This pattern typically reflects social deprivation and a breakdown of conventional socialization, often resulting from incompetent or criminal parents and often in environments of criminal peers or oppositional criminal subcultures, e.g., Lykken's (1995) "common sociopath," Warren's (1971) "subcultural identifier," Jesness's (1988) "socialized conformist," Mealey's (1995) "secondary sociopath" and Miller's (1958) "lower class gang delinquent."
3. *Adolescent-Limited (AL)*: Moffitt (1993) and Lykken (1995) both propose a category of mostly well-socialized normal youth who temporarily associate with or "mimic" delinquent peers. Lykken describes these as normal youth who temporarily engage in a search for meaning, rebellion and sensation seeking, but with a time-limited subcultural identification (Moffitt et al. 2001).

4. *Neurotic or Internalizing Delinquents*: This type is typically characterized by social withdrawal, depression, social anxiety, hostility and mental health problems, often accompanied by severe parental abuse, interpersonal rejection and neglect. Similar types in previous taxometric studies include "Group Baker" (Megargee and Bohn 1979) and "Type D" (Harris and Jones 1999). This general pattern was not included in Moffitt's initial 1993 statement, but she has recently added an analogous internalizing pattern (Moffitt 2003).
5. *Undercontrolled Serious Delinquents—Impulsive and Unsocialized*: This category is generally described as showing early onset of problem behaviors, serious versatile crimes and a personality pattern that is impulsive, risk-taking, aggressive, manipulative, callous, glib and superficially charming. Potential matches include Moffitt's (1993) life-course persistent offenders (LCP); Lykken's (1995) primary psychopath, Mealey's (1995) primary sociopath, Quay's (1990) unsocialized psychopaths and Jesness's (1988) immature aggressive offenders (see also Frick 2004; Hare 1996; Skilling et al. 2001).

Method

Sample

The initial sample consisted of 1,572 juvenile offenders from three jurisdictions. Sixty-three percent of the sample was drawn from the commitment population of a southern state ($n = 999$); 26% from the commitment population of a midwestern state ($n = 414$), and 10% from a probation agency in an urban county of a western state ($n = 159$). The data were collected from 2002 to 2004. The sample is 72% male. The average age is 15.3 years (SD = 1.6) and ranges from 9.2 years to 18.9 years. The ethnicity breakdown in the sample is 54.3% Caucasian, 27.4% African American, 6% Latino/a, 6.4% American Indian, and 5.6% Asian American or other ethnic group. The mean age-at-first-adjudication in the sample is 13.8 years (SD = 1.7); the mean number of adjudications is 2.9 (SD = 2.1); and the mean number of felony adjudications is 1.1 (SD = 1.2). The samples from the three jurisdictions are similar with regard to number of adjudications, but the commitment samples have lower mean age-at-first-adjudication and higher mean number of felony adjudications. The mean age-at-first-adjudication is 13.6 years (SD = 1.5) in the southern commitment sample, 13.9 years (SD = 1.9) in the midwestern commitment sample, and 14.5 years (SD = 1.6) in the probation sample. The mean number of felony adjudications is 1.2 (SD = 1.1) in the southern commitment sample, 1.0 (SD = 1.5) in the midwestern commitment sample, and .5 (SD = .8) in the probation sample.

Measures

The selection of measures in a taxonomic study is of equal importance to any subsequent step (Milligan 1996; Lenzenweger 2004). It will govern explanatory power, interpretative coherence, completeness of type description and ultimately the ability to discriminate between types. Ideally, the classification space should be comprehensive, theoretically guided, coherent and relevant. We aimed to ensure a broad coverage of key features from the prior taxonomic literature. Aside from the taxonomic literature, the conceptual platforms most closely guiding the feature selection were Bronfenbrenner's (1979) ecological

perspective and Farrington's (2003) Integrated Cognitive Antisocial Potential Theory. These identify several domains of risk factors: youth lifestyle and behaviors, youth attitudes and personality, school, leisure activities, peer relations, family, neighborhood characteristics and demographics.

These domains were covered by using the Youth COMPAS (Brennan and Dieterich 2003), a 171-item semi-structured assessment instrument that yields 32 scale scores. Each scale is defined by a set of ordinal level items that have four or five response categories. For example, the response categories on a five-point item range from "definitely no" (1) at one pole to "definitely yes" (5) at the other. Example items from each of the scales are in Appendix A. Table 1 lists the Youth COMPAS scales and provides their respective number of items, mean score, standard deviation and Cronbach's alpha in the estimation sample ($n = 1{,}572$). The average Cronbach's alpha for all scales is .78, and 25 out of 32 scales have alpha coefficients above 0.70. Full details of the Youth COMPAS scales, item content and measurement properties are available from the first author.

Analytic Approach

Our taxonomic procedures follow methods recommended by Skinner (1981) and more recently by Lenzenweger (2004) as follows:

Step 1: Preliminary Transformations

All raw scale scores were first transformed into normalized *Z*-scores with zero mean and unit standard deviation (see Milligan 1996).

*Step 2: Generating Seed Points for K-Means and Preliminary Estimates of the Number of Clusters (*K*)*

We employed a two-stage clustering for preliminary estimates of typological structure. For the first stage, we used Ward's hierarchical method using all 32 scales to analyze three random samples of $n = 300$. A key output of these analyses is an error graph indicating the gradual increase in within-cluster heterogeneity as clusters grow at successive fusions in this hierarchical process. Conventionally, a sudden rise in these graphs suggests a stopping point in the agglomeration process (Aldenderfer and Blashfield 1984). All three Ward error graphs suggested substantial escalation of cluster heterogeneity at $K = 7$ with a further jump at 5. These findings suggest that the taxonomic structure of this population is best reflected by seven subordinate clusters, perhaps nested within five more general clusters. Central prototype cases from these seven clusters were selected for use as starting solutions (seed points) for the next stage using nonhierarchical clustering methods.

Step 3: Examining Taxonomic Structure

Three methods were used to examine taxonomic structure at the seven, six and five cluster levels of the main sample. These were standard K-means, bagged K-means, and semi-supervised learning. This allows cross-method verification of the taxonomic models emerging from each method. These methods are now briefly described with references to relevant expository literature:

Table 1 Summary statistics and reliability coefficients for cluster analysis inputs

Scale	Items	Min.	Max.	Mean	SD	Alpha
Criminal opportunity	6	1	4	2.28	0.67	0.82
Low prosocial	5	1	4	3.09	0.66	0.66
Criminal associates	6	1	4	2.12	0.79	0.82
Social isolation	5	1	5	2.23	1.07	0.87
Common drugs	3	1	5	2.58	1.12	0.72
Hard drugs	7	1	4	1.19	0.37	0.63
Substance trouble	4	1	5	2.44	1.44	0.89
Promiscuity	6	1	5	2.56	0.81	0.66
Impulsivity	5	1	5	3.59	1.05	0.83
Manipulative	5	1	5	3.15	1.01	0.82
Low empathy	4	1	5	2.61	0.98	0.79
Aggression	5	1	5	3.38	1.00	0.81
Violence tolerance	5	1	5	2.81	1.20	0.88
Low remorse	5	1	5	2.44	1.02	0.78
Negative cognitions	5	1	5	2.88	1.08	0.83
Academic failure	3	1	5	2.86	1.06	0.39
Attention problems	4	1	5	3.16	1.20	0.85
Low goals	4	1	5	1.79	1.11	0.77
School behavior	5	1	5	3.23	1.01	0.60
Family discontinuity	5	1	5	3.16	1.19	0.69
Socioeconomic	4	1	5	2.71	1.31	0.88
Family crime drugs	15	1	5	2.37	1.01	0.85
Neighborhood	5	1	5	2.36	1.24	0.88
Parental conflict	4	1	5	2.56	1.35	0.93
Inconsistent discipline	4	1	5	2.33	1.15	0.89
Poor supervision	5	1	5	2.26	1.07	0.80
Emotional bonds	4	1	4	1.89	0.62	0.58
Neglect	4	1	5	2.17	1.06	0.86
Physical abuse	4	1	5	2.04	1.15	0.81
Emotional support	4	1	5	2.26	1.01	0.73
Sexual abuse	4	1	5	1.62	1.02	0.87
Youth rebellion	5	1	5	2.76	1.08	0.79

Classic or Standard K-Means. Classic K-means has been (with Ward's method) a preferred choice in numerous studies to detect latent taxonomic structure (Han and Kamber 2000; Milligan 1996). K-means methods start by creating trial partition—defined by an initial set of seed points or trial cluster centers. K refers to the number of these seed points. The method progressively modifies and refines this initial partition over a series of iterations that attempt to progressively improve a mathematical index of internal cluster homogeneity. Iterations cease when this homogeneity cannot be further improved (for full descriptions, see Aldenderfer and Blashfield 1984; Han and Kamber 2000). We applied classic K-means to the total sample at $K = 5$, 6, and 7 to examine all three hierarchical levels. At each level, pairs of solutions were obtained using (1) optimal seed points as identified above and (2) random seeds.

A weakness of standard K-means is that it is vulnerable to data that do not conform to it's minimum-variance assumption and cluster shapes that may straggle cloud-like in the measurement space. It is also vulnerable to outliers. Thus, we proceeded to use two additional methods designed to address these problems: bagged K-means and the semi-supervised connectivity method.

Bagged K-Means. Bootstrapped aggregation has been used with success for many classification and regression tasks (Breiman 1996). This approach generates multiple bootstrap samples with replacement (bags) from the initial large sample. It then applies K-means to each of these samples to produce numerous classification models. These multiple solutions are then integrated (using a final K-means) to give one final aggregated model. More specifically we used the Bagged K-means implementation in R (R Development Core Team 2006) to generate 1,000 random samples (bags) from our initial sample of 1,572 cases. Each of these was then clustered using standard K-means. The cluster solutions emerging from these were then integrated into a final model by using the cluster centers from these sub-samples as "cases" and re-clustering these with standard K-means. To check the stability of this analysis we used both randomly initialized seed points and also seeds from the initial K-means analysis (Step 2: Generating Seed Points for K-Means and Preliminary Estimates of the Number of Clusters (*K*)). These converged on essentially the same cluster solution. This resulting cluster model was then used to provide centers for an overall K-means on the total sample to create a final taxonomic model.

Semi-Supervised Clustering. Zhou et al. (2004) introduced the consistency method, a semi-supervised clustering technique. Given the recent emergence of this approach its mathematical details are given in the Appendix, with basic expository references. This approach discovers clusters by allowing them to grow from initial starting seeds—according to a criterion of internal connectivity among members. It is referred to as "semi-supervised" because it requires known members (usually seed points for the suspected class) to start this growth process. Cluster growth continues until connectivity weakens at cluster boundaries. Whereas K-means tends to favor (or impose) hyper-spherical clustering shapes, the semi-supervised method may detect arbitrary cluster shapes constrained only by internal connectivity. It is thus more likely to discover "natural clusters" (Arabie et al. 1996; Gordon 1999).

Specifically, we applied the semi-supervised method to our initial sample of 1,572 cases. For "known members" we used the seven cluster centers previously established from the final bagged K-means to initiate the process. A useful feature of this method is that multiple trials can be conducted by systematically varying the connectivity parameters to create several alternative semi-structured cluster solutions. This offers an additional internal reliability analysis (see Appendix B for mathematical details).

Step 4: Identifying Core Prototypes—Producing a Consensus Typology

This step creates a refined or consensus typology consisting only of those cases consistently assigned to the same clusters across the main methods. To do this we first identified matched cluster pairs by comparing the cluster models from the final semi-supervised method and final bagged K-means model (both from Step 3) and noted the highest percentage overlap between cluster pairs from the two solutions. Consistently classified cases are those that were jointly assigned to the same matched pairs by the two methods. The remaining inconsistently classified cases are likely to be closer to the boundaries of more weakly classified cases and are likely to be outliers or hybrids located between clusters. The two cluster methods essentially cast "votes" for the cluster assignment of each case.

The refined solution is simply attained by eliminating all inconsistently classified cases and should contain only the more central core members of the seven types. We refer to this as the "consensus" model.

Results

Estimating the Number of Clusters (K)

Preliminary tests with Ward's hierarchical clustering of the three sub-samples with the pseudo-F index (Calinski and Harabasz 1974) and the pseudo-t squared method of Duda and Hart (1973) were inconclusive, suggesting that several levels between 2 and 7 might be appropriate. However, all three error-graphs in the Ward analyses started trending upwards at $K = 7$, with a further break at $K = 5$ suggesting taxonomic structure at these two levels, see Milligan and Cooper (1985) and Aldenderfer and Blashfield (1984). Additionally, the substantive interpretations of K-means partitions at $K = 5$, 6, and 7 levels indicated that the 7-level was both highly interpretable and easily related to the 5-level. Thus, $K = 7$ was chosen as a preferred level of the hierarchical taxonomic structure.

Convergence Across Clustering Methods

Cohen's kappa coefficient was used to compare all pairs of classifications emerging from these analyses. As a measure of agreement between classifications it is generally seen as more appropriate than a simple percent agreement since it takes account of agreements due to chance. In comparing the partitions of the 7-level from the standard K-means and bagged K-means, a kappa of 0.65 indicated that the two methods were converging on largely overlapping high density regions of the space. The cluster interpretations from these two solutions were also very similar, although this kappa level suggests that the boundaries are not identical.

In comparing the several semi-supervised method partitions to the basic and bagged K-means, we found lower matching, yet with still encouraging agreement (kappa varied between 0.555 and 0.583 for each pair of solutions). This slightly lower kappa stems from the fact that K-means and semi-supervised methods generate clusters in different ways with different shape constraints. Surprisingly, even at these lower kappa levels, the interpretations of the seven clusters were similar across solutions. This suggests that while boundary conditions of clusters may be unstable, the central exemplars or core members of the clusters appear relatively stable and may largely determine cluster interpretation (see Cluster Descriptions below).

Refinement of Clusters (Voting Procedure): The Core Consensus Model

The refined consensus model that retains only the consistently classified core members of each cluster showed considerable agreement with the previous main models (standard K-means, bagged K-means and semi-supervised models). Specifically, it showed strong agreement ($\kappa = 0.949$) with the initial K-means-only solution and the bagged K-means solution ($\kappa = 0.992$). Such strong matching was not unexpected given that all outliers and inconsistently classified cases had been removed from this consensus model. We report the interpretation of this core consensus model in the cluster profile descriptions (below), although we found virtually identical profiles from the full sample bagged K-means model.

Examining the refined consensus model can indicate the proportion of the sample that was consistently clustered or that remained unclassified. Table 2 shows that 57% of the sample was consistently classified as "strongly clustered" under this voting rule and may represent the core or central cases of each cluster. Conversely, 43% remained unclassified and are likely to be hybrids or outliers. This finding is tentative, and the percentage of unclassified cases should be treated with caution. It could change if a different voting procedure were used. The present voting procedure may be quite stringent since it requires joint matching across two clustering methods that detect clusters in quite different ways and which impose different boundary conditions and "shapes" on the clusters.

External Validation: Cross-Classification Against Criminal History and Demographics

External validation examines whether there are significant differences between the seven consensus model clusters on several external (but relevant) variables not included in cluster development. Table 3 shows the results of post-hoc tests (Dunnett's C) comparing the cluster means for the following external variables: age-at-assessment, age-at-first-adjudication, total adjudications, and total violent felony adjudications. These results show several significant differences—all in expected directions. For example, clusters 4 and 7, which both partially match the low-risk profile of Moffitt's (1993) AL type, have

Table 2 Prevalence of well-classified cases and hybrids from the consensus model

Cluster	Frequency	Percent
Hybrid cases	677	43.1
Internalizing A	83	5.3
Socially deprived	103	6.6
Low control A	85	5.4
Normal accidental	151	9.6
Internalizing B	197	12.5
Low control B	146	9.3
Normative delinquents	130	8.3
Total	1572	100.0

Table 3 Means (standard deviations) of age-at-assessment and criminal history measures by cluster

External indicator	Cluster							F*
	1	2	3	4	5	6	7	
Age-at-assessment	14.93[a] (1.83)	15.17[b] (1.68)	15.68[c] (1.51)	15.37[d] (1.57)	14.50[cde] (1.57)	15.80[ae] (1.62)	16.08[abd] (1.53)	17.15
Age-at-first adjud.	13.48[a] (1.65)	13.70[b] (1.63)	13.93 (1.53)	14.19[de] 1.56)	13.44[df] (1.61)	13.36[eg] (1.69)	14.47[abfg] (1.64)	9.31
Total adjudications	2.52[ag] (1.71)	2.98 (1.93)	3.62[abc] (2.58)	2.46[bd] (1.88)	2.54[cf] (1.42)	3.59[dfg] (2.54)	2.92 (1.65)	7.15
Violent felony adjud.	.25 (.46)	.25 (.52)	.16[c] (.43)	.31 (.56)	.46[ce] (.66)	.30 (.61)	.18[e] (.46)	4.75

Note. Means with the same superscript are significantly different from each other (Dunnett's C post-hoc tests)

* All main effects are significant at $p < .001$

significantly later age-at-first-adjudication compared to the high-risk cluster 6 that matches Moffitt's high-risk LCP and Lykken's (1995) psychopath category. Cluster 6 in addition has the earliest age-at-first-adjudication and significantly higher total adjudications than clusters 1, 4, and 5—again consistent with Moffitt's descriptions. Other significant differences are discussed in the cluster interpretation section below.

Cluster Descriptions

This section describes the seven empirical patterns from the consensus solution. The total sample bagged K-means and classic K-means produced essentially similar substantive descriptions and it would be redundant to also present these. Theoretical interpretations and links to prior taxonomies are delayed until the discussion section.

Figure 1 provides a bar chart of each feature for each cluster in the core consensus solution. A bar extending to the left indicates a cluster has a lower mean score than the grand mean. If the bar extends to the right the cluster has a higher mean score than the grand mean. The scale inputs in the bar chart are sorted by domains representing peers, substance abuse, attitudes, cognitions, school, neighborhood and family. The bar chart was created using a modified function in the pamr package (Hastie et al. 2006) in R (R Development Core Team 2006). The numerical Z values of the seven cluster centroids are provided in the following descriptions (in parentheses). A complete table of the numerical Z values is available from the first author.

Cluster 1. Internalizing Youth A: Withdrawn, Abused and Rejected

This cluster reflects social isolation (.78), negative social cognitions and mistrust (.70) and an aggressive hostile attitude to others (.42). They appear to follow a low-risk lifestyle (−.59) and avoid criminal peers (−.57), drugs (−.70) and promiscuity (−.22). However, they experience extreme physical abuse (.98), sexual abuse (.53), emotional rejection (.71) and neglect (1.04), as well as weak discipline (.98) and supervision (.66). Their families are of lower socioeconomic status (.74), disorganized (.62) and above average for crime/drugs (.33). The results in Table 3 indicate that this cluster has significantly earlier mean age-at-first-adjudication than cluster 7, but significantly fewer mean adjudications than clusters 3 and 6.

Cluster 2. Socially Deprived: Subcultural or Socialized Delinquents

This cluster exhibits social deprivation and inadequate socialization. Their families reflect lower socioeconomic status and poverty (.42), higher parental crime/drugs (.53), family disorganization (.49) and parental conflict (.44). The youth have poor emotional support at home (.32), inconsistent discipline (.23) and poor school attainment (.30). Yet, interpersonally, they are not socially isolated (−.41), show little hostile negative social attributions (−.44) or aggression towards others (−.44) and few low-control personality features. This cluster is about average for total adjudications and age-at-first-adjudication. It has a significantly earlier mean age-at-first-adjudication compared with cluster 7.

Cluster 3. Low-Control A: Versatile Offenders

This cluster (with Cluster 6) reflects low self-control, i.e., impulsivity (.72), low empathy (.68), manipulative-dominance (.75), aggression (.84) and low remorse (.62). Their

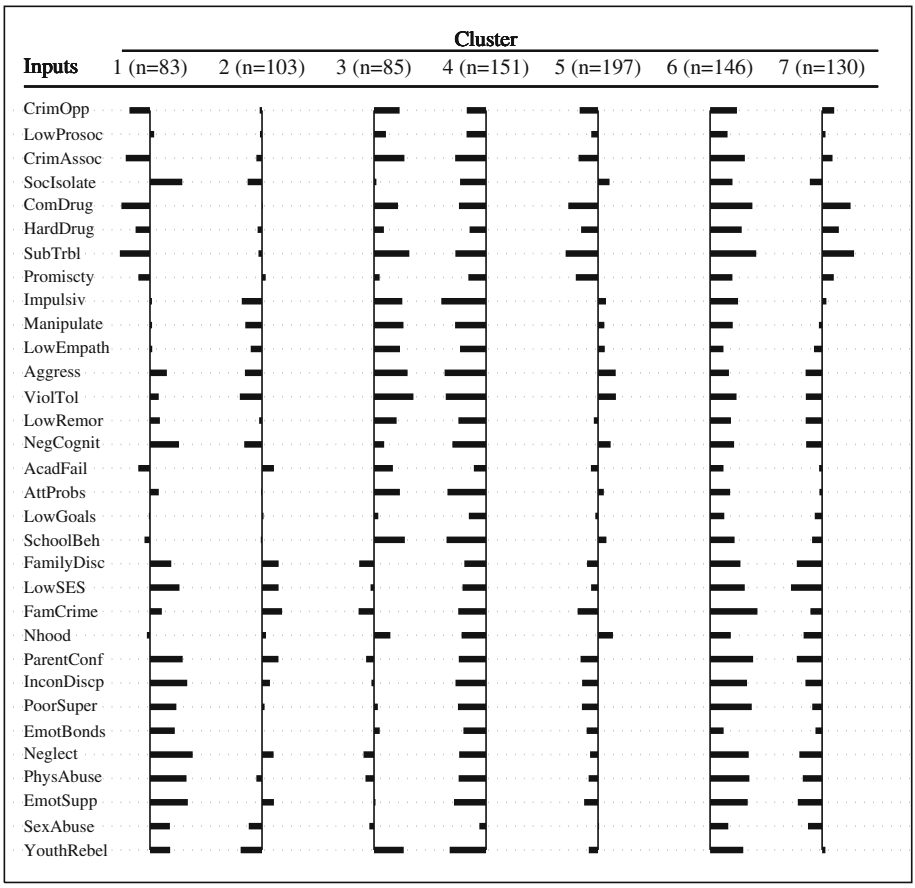

Fig. 1 Bar chart of cluster centroids from the consensus solution

lifestyle involves criminal peers (.84), high criminal opportunity (.69), common drugs (.60) and unsafe neighborhoods (.38). Their school experience is negative: school failure (.48), attention problems (.64) and disruptive school behavior (.78). While parenting features are not extreme, these youth are in open rebellion against parents (.75). This cluster's official criminal history coheres with the above profile. This cluster has a significantly higher mean number of adjudications than clusters 1, 4, and 5.

Cluster 4. Normal "Accidental/Situational" Delinquents

This cluster exhibits few risk factors and a benign personality that is consistent with their low official delinquency and later onset, i.e., stable families (−.51), safer areas (−.67), good supervision (−.71), good discipline (−.76), fewer high crime peers (−.74) and lower-risk lifestyles (−.48). There are no obvious problems at school. These youth also mostly avoid drugs (−.42) and promiscuity (−.38). They are not socially isolated (−.70). The external official criminal record coheres with this benign profile. This cluster has significantly later mean age-at-first-adjudication compared with clusters 5 and 6 and significantly lower mean number of adjudications compared with clusters 3 and 6.

Cluster 5. Internalizing Youth B: With Positive Parenting

This cluster reflects an internalizing pattern (like Cluster 1) with above average negative cognitive attributions and mistrust (.30), hostile aggression (.44) and above average social isolation (.25). Like Cluster 1 they follow a relatively low-risk lifestyle (−.46), avoiding delinquent peers (−.44), common drugs (−.73), hard drugs (−.43) and promiscuity (−.49). However, they differ from Cluster 1 by the presence of caring, competent and non-abusive parents, who are not emotionally rejecting (−.33) or neglectful (−.22) and who provide fair supervision (−.41) and discipline (−.39). These families appear less disorganized (−.23) with lower family crime/drugs (−.48) and lower parental conflict (−.42) than average. The results in Table 3 indicate significantly younger mean age than clusters 3, 4, 5, and 6; significantly earlier mean age-at-first-adjudication than clusters 4 and 7; significantly higher mean number of violent felony adjudications than clusters 3 and 7; but significantly lower mean number of adjudications than clusters 3 and 6.

Cluster 6. Low-control B: Early Onset, Versatile Offenders with Multiple Risk Factors

This low control, high delinquency type is a more extreme variant of Cluster 3. It shows multiple antisocial personality features: impulsivity (.71), low remorse (.58), low empathy (.36), aggression/hostility (.47) and manipulation (.58). It has the worst parenting, with very low emotional support (.96), neglect (.93), weak supervision (1.04), inconsistent discipline (.97) and physical abuse (1.05). These families also appear seriously conflicted (1.11) and disorganized (.85), with high criminality/drugs (1.21) and residence in unsafe areas (.49). These youth have delinquent peers (.96), high-risk lifestyles (.73), few prosocial activities (.46), common drug use (1.06), hard drug use (.79) and promiscuity (.63). At school they report attention problems (.49), behavioral problems (.62) and low academic goals (.44).

This type is more extreme than Cluster 3, with higher scores for drugs, promiscuity, abuse, neglect and parent criminality. They also present the internalizing pattern of social isolation (.53), aggressive hostility and negative cognitive attributions (.58). The external official history variables cohere strongly with this extreme pattern. The results in Table 3 indicate significantly earlier mean age-at-first-adjudication than clusters 4 and 7, and significantly higher mean number of adjudications than clusters 1, 4, and 5.

Cluster 7. Normative Delinquency: Drugs, Sex and Peers

This cluster (like Cluster 4) reflects more "normal" youth with few risk factors. Their families are not from a lower socioeconomic strata (−.78), they live in safer areas (−.52), have relatively good family stability (−.60), are not abusive (−.47) and provide consistent discipline (−.40). At school there are few signs of failure or behavioral disruption. Personality has few signs of callous lack of remorse (−.37), aggressive hostility (−.43), tolerance for violence (−.44), social isolation (−.35) or negative cognitive mistrust (−.40). However, compared to Cluster 4, they have higher scores for antisocial peers (.33), risky lifestyle (.34), promiscuity (.36), conventional drugs (.71) and hard drugs (.42). Thus, their risks (drugs, sex, and peers) occur in a context of several strengths. The results in Table 3 indicate significantly later mean age-at-first-adjudication than clusters 1, 2, 5, and 6, and a significantly lower mean number of adjudications than clusters 1, 2, 4, 5, and 6.

Cross Replication on an Independent Sample

Replication Sample

A replication sample (B) of 1,453 youth was assessed using identical instruments. This sample consisted of successive admissions to juvenile assessment centers at four urban judicial districts in a western state that was not included in the training sample. The sample is 67% male. The average age is 15.6 years (SD = 1.6) and ranges from 9.0 years to 18.0 years. The ethnicity breakdown in the sample is 54.2% Caucasian, 17.5% African American, 23.9% Latino/a, and 4.4% other ethnic groups. Approximately 70% of these youth had entered the juvenile assessment centers after an arrest for a misdemeanor or felony offense, while the remainder were brought in for others reasons, including status offenses, school referrals, and family issues. Fifty-five percent of the sample had no adjudications. Sample B contains fewer serious delinquents than the original sample A.

Cross-Replication Method

We used the McIntyre and Blashfield (1980) cross-replication design, specifically for cluster analysis, to assess whether similar types emerge across the two samples. It has three steps:

1. Repeat the original analysis on the replication sample to provide an independent typological structure (Partition B1). We again used bagged K-means with 1,000 replications since this approach classified ALL cases and was essentially similar to the consensus solution (which does not contain ALL cases).
2. Mimic the original cluster partition using the replication sample. This step assesses whether the replication sample B cases can meet membership criteria for the clusters of the original sample cluster model. This step uses a pattern matching procedure—a Support Vector Machine (SVM) - that was trained as a classifier using the initial bagged K-means 7 cluster model from sample A. SVM methods are designed to classify unknown cases into any given prior classification. It proceeds by building optimally separating boundaries around the given classes (known as "maximum margin hyperplanes"). A case can only be assigned to a given class if its profile vector falls inside the maximally separating boundary for that class. The method has performed well in comparative studies against other classifiers. Detailed descriptions are given in Han and Kamber (2000). The trained SVM method was applied to the replication sample to identify any cases in sample B that successfully met the boundary conditions of the original model clusters from sample A. Cases not meeting the SVM boundary criteria would remain unclassified. This gives partition B2.
3. Compare partition B1 to partition B2. This tests whether the independent replication cluster model (B1) matches partition B2 (from the SVM) that mimics the original taxonomic model. This analysis also estimates the percentage of cases in the replication sample that remain unclassified.

Cross-Replication Results

1. Agreement between initial and replication models: Partition (B2) and partition (B1) showed very strong and significant global similarity (Contingency Coefficient = 0.84, $p > .0001$). The contingency table indicated that this strong association resulted from the fact that about 60% of cases were jointly classified into the identical matched

cluster pairs across the two partitions. However, only six of the original seven patterns were substantially rediscovered in the replication sample by the independent clustering (B1). The initial cluster 6, the most extreme delinquent cluster, was not recovered in the clustering of the replication sample. This result appears due to the rarity of cluster 6 in the replication sample. The SVM findings indicated that, in fact, a small number of cluster six caseswere found in the replication sample: 1.6% for boys and 1.9% for girls—seeTable 4. This result is understandable because the most serious delinquentswould be unlikely to be referred to the juvenile assessment centers.

2. Re-emergence of specific cluster patterns: In an elaboration of the McIntyre-Blashfield method we ran five additional cluster analyses on the replication sample using the same standard and bagged K-means but with different seed points. We also ran the semi-structured consistency method. In these additional analyses we asked for only six clusters (since cluster 6 was basically missing). These five additional models—plus the original SVM consensus model— give six separate estimations of each of the six replicated cluster profiles. Regarding these as multiple measures of each of the six replicated clusters, we computed the average correlation of each original cluster to its five replications in the independent sample. Table 4 shows the replication statistics for each original pattern. The average correlation of each original cluster with its five replications ranged from .93 to .45, suggesting excellent to substantial replications.
3. Cluster percentages in the replication sample: The SVM indicated that overall 70% of the replication sample was assigned to one of the original seven patterns (Table 4). The most frequent cluster for both boys and girls was the relatively low-risk/low-delinquency cluster 4. Cluster 6 (the most serious delinquent profile) was the least frequent, with only 1.6% for boys and 1.9% for girls. Additionally, all seven of the original clusters were recovered by the SVM; although, as noted, very few cases in the replication sample matched cluster 6. Overall, 30% of the replication sample failed to meet the matching criteria of the SVM and remained unclassified.

Discussion

Do Clusters Exist?

The debate over the "reality" of delinquent types in the delinquency literature rarely discusses the structural aspects of the proposed taxonomies (Rosch 1978; Bryant 2000) but

Table 4 Replication statistics for each initial cluster pattern

	Patterns from the initial sample						
	1	2	3	4	5	6	7
Average correlation of each initial pattern with 5 replicated patterns	.87	.93	.83	.52	.45	NA	.92
SVM: % of Boys in replication sample classified to initial clusters[a]	3.40	7.40	8.50	24.90	13.30	1.60	10.00
SVM: % of Girls in replication sample classified to initial clusters[b]	6.30	9.10	6.30	22.70	13.30	1.90	13.20

[a] 30.0% unclassified

[b] 31.4% unclassified

usually focuses only on substantive interpretations and commonalities across studies. This lack of concern for taxonomic structure suggests that different participants in this debate may have quite different ideas on what constitutes a distinct type and different ideas on what structural criteria to use. Thus, before examining continuities with the prior delinquent taxonomies of Moffitt, Lykken and others, we briefly review the structural aspects of the taxonomy emerging in this research.

Internal Structures—Stable Cores and Graded Membership

The internal structure of our offender classes strongly suggests family resemblance clustering or Roschian polythetic categories (Rosch 1978; Bryant 2000). Polythetic classes are characterized by fairly stable, relatively dense cores of mutually similar cases (exemplar or prototype cases), with less typical cases shading off probabilistically toward poor-fitting cases at cluster boundaries. Polythetic categories also avoid reliance on a fixed set of necessary and sufficient factors as required by monothetic (classical) categories (Brewer 1993).

The methods we used (Ward, complete link, K-means, bagged K-means) are all density-seeking in the multivariate space. Their recurring convergence on similar clusters suggests that these dense regions may represent stable reference points reflecting dominant patterns. The kappa coefficients between our trial classifications that used all cases ranged from 0.55 to over 0.70, which suggests substantial convergence (see Costa et al. 2002). Particularly noteworthy was the agreement between the bagged K-means and initial standard K-means (κ = 0.949), because these analyses preceded removal of weakly classified cases and utilized little parameter tuning, except for the number of clusters (K). The substantive interpretations of these dense regions were also relatively unchanged before and after the removal of peripheral cases. Kappa coefficients of over .90, following the voting procedure and the identification of the consensus core members, indicated the almost perfect stability of these refined core clusters, irrespective of clustering method. These findings underline a point noted by Brewer (1993) that natural classes are largely defined not by the boundary lines, but by the central cases; not by what is excluded, but by what they most strongly include. This consistency across methods offers evidence of the stability of the core pattern structures. Thus, we conclude that the distribution of delinquents in this measurement space is non-random. Certain regions are relatively dense and stable, while other regions are relatively sparse.

Unstable or Fuzzy Boundaries, Hybrids and Unclassifiable Cases

Our analyses also show that cluster boundaries are relatively unstable. Kappa coefficients from 0.55 to 0.70, although indicating general overlap, also imply that boundaries between clusters may be imposed differently by different analyses, and cases close to boundaries may be unreliably classified across adjacent clusters. Such cases may be regarded as hybrids. Lykken (1995) has argued that even an ideal taxonomy may have borderline or hybrid cases (p. 21).

Based on the results from the voting procedure, we initially estimated that 43% of the sample may be regarded as weakly classified or hybrid cases. But the voting criterion of "joint agreement" between two methods of clustering that differ in terms of cluster definitions, ways of identifying clusters, and approaches to boundary and shape constraints, may be a fairly stringent requirement. When the SVM procedure was applied in the replication analysis we found that about 70% of the sample could be consistently classified,

leaving only 30% unclassified. Thus, the percentage of unclassifiable cases must be tentative. There are many other approaches to identifying borderline or outlying cases (Bergman 1988; Wishart 2003).

The Hierarchical Structure—from Global Superordinate to Fine-Grained Subclasses

Our findings also imply a hierarchical taxonomic organization with five more inclusive (superordinate) classes and seven more fine-grained subordinate classes. This is not surprising. Warren (1971) presented a hierarchical structuring in the original I-Level system with several fine-grained subtypes nested within broad moral developmental categories. Similarly, Lykken's (1995) three broad "genera" (sociopaths, psychopaths and neurotic types) contain numerous finer grained subcategories. We selected the $K = 7$ level for descriptive interpretation and tentatively offer this as a basic taxonomic level. Rosch (1978) describes the "basic level" of a hierarchical taxonomy as that particular cut or partition that captures the most useful distinctions in the fewest categories. Our choice of $K = 7$ must also be tentative since current methodological approaches to choosing an optimal hierarchical level of a taxonomy remain controversial (Milligan and Cooper 1985; Raftery and Dean 2004).

Replications and Extensions of Prior Taxonomies

We now offer conclusions regarding replications and extensions of prior delinquent taxonomies. Unfortunately, most prior theoretical and quantitative studies omitted operational procedures for type matching (such as SVM). Thus, our comments are offered only at a logical or descriptive level by examining commonalities of profile elements.

Cluster 1. Internalizing Youth A: Withdrawn, Abused and Rejected

A pattern of parental abuse linked to the internalizing features of social withdrawal, hostility and suspicion has emerged in several prior studies. Lykken (1995) described a "neurotic" type exhibiting social withdrawal, suspicion, inadequate personality and low self-esteem in a context of rejection and uncaring relationships. Similar types include Miller et al.'s (2004) "internalizing depressed", Aalsma and Lapsley's (2001) "Cluster 2" and Harris and Jones's (1999) "type E." Moffitt's (2003) recent work—using Ward's clustering—also adds an internalizing group with substantial psychological problems. Others suggest that this internalizing pattern may involve trauma from family abuse that impairs parent-child attachments and undermines bonding abilities (Miller et al. 2004; Raine et al. 2005).

Cluster 2. Socially Deprived: Subcultural or Socialized Delinquents

The social deprivation of cluster 2 appears to replicate the "lower class" or "socialized" delinquent often described in the early sociological literature (Jesness 1988; Miller 1958; Van Voorhis 1994; Warren 1971). The features of cluster 2 (poverty, criminal/drug-using parents, family disorganization, poor discipline, neglect and school failure) also match Lykken's (1995) "common sociopath," described as poorly socialized but psychologically "normal." Cluster 2 also is consistent with Lykken's description of a broadly normal temperament, because it shows little evidence of either low social control or the internalizing pattern. This cluster pattern does not appear in Moffitt's theoretical taxonomy.

Cluster 3. Low-Control A: Versatile Offenders

The present analysis, in contrast to Moffitt, but again consistent with Lykken, produces two impulsive, potentially psychopathic clusters (3 and 6). These largely overlap with Lykken's primary and secondary psychopaths, although we acknowledge this distinction remains controversial (Newman et al. 2005). Cluster 3 shares several features with Lykken's (1995) primary psychopath, including impulsivity, low empathy, hostility, manipulative-dominance, low remorse, attention problems, disruptive school behaviors, criminal peers, high-risk lifestyle, drug abuse and serious criminal history. Other close replications include Moffitt's LCP, Alterman et al.'s (1998) "Cluster 5 Psychopathic" and Vincent et al.'s (2003) "Impulsive" cluster.

Two features of this cluster may address prior theoretical controversies. First, this cluster has no evidence of the socially withdrawn internalizing pattern—which matches Lykken's confident and domineering primary psychopath versus his introverted secondary psychopath (see also Blackburn 1986). Secondly, the lack of evidence for dysfunctional parenting supports the profiles of Lykken (1995) and Alterman et al. (1998) and offers a challenge to Gottfredson and Hirschi's (1990) assertion of a general link from inadequate parenting to low self-control (Cluster 6, however, clearly supports their assertion of a link from dysfunctional parenting to low control).

Cluster 4. Normal "Accidental/Situational" Delinquents

Cluster 4 reflects mostly low-delinquency youth with very few risk factors. This benign pattern, with a relatively late age-at-first-adjudication and mostly minor delinquency, appears consistent with Moffitt's AL type. Lykken (1995) similarly describes a type with a normal temperament, reasonably good socialization and competent parents. Other studies with a similar "normal" type include Simourd et al. (1994), Aalsma and Lapsley (2001), Harris and Jones (1999) and Huizinga et al. (1991). Delinquency in such youth is commonly explained by situational or accidental factors or peer influence (Warren 1971; Lykken 1995; Van Voorhis 1994).

Cluster 5. Internalizing Youth B: with Positive Parenting

Clusters 5 and 1 both exhibit an internalizing pattern of social withdrawal, isolation and mistrust. Both avoid delinquent peers, drugs and sex and have low adjudication rates. However, in contrast to the abusive, violent parents of cluster 1, the parents of cluster 5 appear mostly non-abusive, competent and caring. Another surprising difference is that cluster 5 has a higher violent felony adjudication rate.

Cluster 5 appears to be a close replicate of one subtype within Lykken's (1995) superordinate "neurotic" category that he describes as having positive parenting and normal socialization—but who are beset by some unconscious or emotional complexity. Harris and Jones (1999) also identified a similar "internally conflicted" cluster. The emergence of both clusters 1 and 5 raises the possibility that at least two different causal processes may underlie this broad internalizing pattern.

Cluster 6. Low-Control B: Early Onset, Chronic Versatile Offenders with Multiple Risk Factors

Cluster 6 is a more extreme undercontrolled variant of Cluster 3. They exhibit drug use, promiscuity, criminal peers, multiple school attention problems and the highest parental

crime and parental abuse. Their official record shows the earliest age-at-first-adjudication and the highest total adjudications. This cluster has features in common with Lykken's secondary psychopath (1995), Mealey's (1995) primary sociopath and Moffitt's LCP category. Other close matches include Sorenson and Johnson's (1996) "Cluster 3 Distressed;" Blackburn's (1995) "Secondary Psychopath;" Aalsma and Lapsley's (2001) "Cluster 3 psychopath" and Alterman et al.'s (1998) "Psychopathic." The finding of two variants of this undercontrolled type challenges Moffitt's taxonomy while supporting Lykken's distinction of primary and secondary psychopathic types.

Cluster 7. Normative Delinquency: Drugs, Sex and Peers

This cluster—with cluster 4—again reflects more "normal" youth with substantial school and family strengths. Unlike Cluster 4, these youth have vulnerabilities to drugs, sex and peers. Their official record shows a later age-at-first-adjudication and mostly non-violent offenses (Moffitt et al. 2001).

Several studies have identified a similar cluster. Most notably it strongly reflects Moffitt's (1993) AL category. Lykken (1995) also described a "dissocial sociopath" type as psychologically "normal" youth engaged in a search for meaning and excitement that may involve drugs and sex. Other similar clusters include Harris and Jones' (1999) "average normal—cluster 8" and Alterman et al.'s (1998) "drug-only clusters."

In conclusion, these results appear to corroborate certain prior categories and extend the taxonomic proposals of Lykken (1995), Moffitt (1993) and Mealey (1995). They also confirm several "consensus" types from the earlier literature on delinquent taxonomies (Warren 1971; Van Voorhis 1994). The following points may be noted:

First, we have added empirical precision and detail to several prior theoretical profiles that have only been inconsistently identified in prior taxonomic studies. The present cluster descriptions may help characterize these delinquent types and lead to explanatory understanding of these different pathways.

Second, our replication study on a large independent sample supports the generalizability of six of the seven types. The failure of cluster 6 to replicate is perhaps explained by the lower delinquency of the replication sample and the rarity of the most serious delinquent cluster 6. However, a few cases of this type were identified in the replication sample by the SVM procedure.

Third, we briefly examined the "number of types" question raised by Moffitt (2003). Most prior taxonomic studies suggest from 2 to 5 subtypes—often with only limited quantitative examination of this issue (But see Harris and Jones 1999; Stefurak et al. 2004). We tentatively suggest that the most likely basic level for an explanatory taxonomy of delinquency consists of either five or seven classes. This would imply that the 2-class taxonomies of both Moffitt (1993) and Mealey (1995)—while identifying important superordinate classes—may be oversimplified and that a more fine-grained hierarchical taxonomy may be required to address the diversity of the explanatory patterns that underly delinquency.

Fourth, our results suggest that while the boundaries of the seven identified clusters may be unreliable and fuzzy, the core members of prototypes of each cluster were sufficiently stable that they could survive both cross-method and cross-sample replication tests. These findings challenge Osgood's (2005) preference (and most theoretical criminologists') for a dimensional structure by offering empirical support for the existence of some degree of taxonomic structure in a measurement space of explanatory factors. On the other hand, and in defense of Osgood's position, we acknowledge that the taxonomic position is weakened

by our demonstration of only fuzzy boundaries between clusters and a substantial percentage (30–40%) of apparently unclassified cases that may represent noisy hybrids and outliers.

Limitations

The present research has several limitations. The initial sample consisted of successive intakes at youth corrections agencies in two states and a county probation agency in a third state. Although large and fairly heterogeneous, the sample did not cover the entire spectrum of juvenile justice agencies, which limits the generalizability of our findings. However, the replication sample consisted of successive intakes at juvenile assessment centers in four large urban judicial districts in a fourth state and was organizationally distinct from the initial sample. Yet, six of the seven clusters re-emerged in the replication sample, which provides support for the validity of the initial findings.

The selection of taxonomic methods is a difficult issue since many alternative approaches are possible. Nagin and Paternoster (2000) noted there is no consensus on the most appropriate methods to study population heterogeneity and suggest that researchers explore different methods with different assumptions. Alternatives include several families of cluster analysis, latent class models, Meehl and Yonce's (1994) taxometric methods and semi-parametric mixed Poisson models (Brennan 1987; Milligan 1996; Nagin and Paternoster 2000). We adopted this mixed approach by using both classical density-seeking methods and a more recent method based on internal connectivity.

Another limitation is the unresolved challenge of finding an optimal value of K. Milligan and Cooper (1985) list over 30 approaches to this problem. Ultimately, as in several recent studies (Costa et al. 2002), we relied on a combination of methods as well as interpretative clarity. Our selection of the $K = 7$ solution is thus tentative.

In conclusion, we agree with Nagin and Paternoster (2000) and Lykken (1995) that we are at early stages in mapping the taxonomic heterogeneity of delinquency—from both behavioral and explanatory perspectives.

Appendix A

Youth COMPAS

The following briefly describes the Youth COMPAS scales. Scale name abbreviations and factor loadings (in parenthesis) for selected key items in each scale are indicated. Full psychometric characteristics and theoretical justifications for each scale are available from the first author.

Antisocial Opportunity (CrimOpp): Hang around with friends (.66), parties w/o adults (.69).

Absence of prosocial engagement (LowProsoc): Church activities (.66), sports, music/hobbies (.66), school activities (.70).

Anti-social peers (CrimAssoc): Friends use drugs (.75), friends arrested (.70), friends dropped out (.65).

Social isolation (SocIsolate): Has trouble making friends (.77), no close friends (.62)

Common drugs (ComDrug): Alcohol use (.81), marijuana use (.78).

Hard drugs (HardDrug): Used cocaine (.72), used heroin (.62), has injected (.48).

Substance abuse trouble (SubTrbl): Poor judgment when high (.87), violent feelings when high (.83)

Sexual promiscuity (Promiscty): Frequency of intercourse (.78), number of partners (.68).

Impulsivity (Impulsiv): Takes risks (.64), makes quick decisions (.76), seen as reckless (.73).

Manipulative-dominance (Manipulate): Good at talking one's way out of trouble (.67), easily lies and gets away with things (.73), can dominate/threaten others (.55).

Empathy (LowEmpath): Feels sad when seeing other people cry (.81), guilt feelings when breaking a promise (.84).

Aggression/anger (Agress): Quick temper (.79), History of fights (.66), stays calm in arguments (.53).

Tolerance of violence(ViolTol): How wrong it is to hit someone to win an argument (.50), hit someone to teach them a lesson (.76).

Lack of remorse (LowRemor): Blames others/situation (.72), doesn't express regret (.63).

Negative social cognitions (NegCognit): Kids put you down (.78), few kids can be trusted (.63).

Academic failure/success (AcadFail): Usual grades (.79), number of classes failed (.80), times grade repeated (.57).

Attention problems (AttProbs): Trouble paying attention (.85), easily bored (.71).

Educational aspirations (LowGoals): Intends to graduate (.72), education is important (.81).

School behavior (SchoolBeh): Suspended (.72), argues/fights with students (.67), conflict w. teachers (.74).

Family discontinuity (FamDisc): Multiple caretakers (.64), separated from natural parent (.61), out-of-home placements (.72).

Social class and poverty (LowSES): Family receives social assistance/subsidized housing (.68), parent has unstable/low wage employment (.81), difficulty paying bills (.87).

Family criminality (FamCrime): Mother arrested (.62), father jailed (.60), sibling drug use (.51).

High crime neighborhood (Nhood): Friends or family assaulted (.73), drug sales (.84), witnessed fights/gunfire (.87).

Parental conflict/domestic violence (ParentConf): Parents threaten each other (.84), parents yell/fight (.78), parents attack each other (.82).

Inconsistent discipline (InconDiscp): parents have clear rules (.72), perceived fairness (.72), clear reasons for punishments (.75).

Inadequate supervision (PoorSuper): Parents check when youth returns home (.84), parents check on youth's friends (.65), parents monitor youth's activities (.85).

Emotional bonding with parents (EmotBonds): Feels close to mother (.73), close to father (.51), feels close to sibling (.70).

Parental neglect (Neglect): Youth feels neglected (.80), parents show no interest (.71), parents rarely talk to youth (.79).

Physical abuse (PhysAbuse): Youth is scared of being hurt (.86), parents violent when high/drunk (.78), youth removed from home because of abuse (.73).

Emotional support (EmotSupp): Mother is hostile (.64), kicked out of the house (.50).

Sexual abuse (SexAbuse): Sexually abused by family member (.81), removed from home because of sexual abuse (.74).

Youth rebellion (YouthRebel): Youth intimidates parent/caretakers (.65), youth openly defies parents/caretakers (.82).

Appendix B

Algorithm

Given a set of points $X \in \mathcal{R}^{n\times m}$ and labels $\mathcal{L} = \{1, \cdots, c\}$. Let x_i denote the ith example. Without loss of generality the first l points (1 ...l) are labeled and the remaining points ($l + 1$ ...n) unlabeled. Define $Y \in \mathcal{N}^{n\times c}$ with $Y_{ij} = 1$ if point x_i has label j and 0 otherwise. Let $\mathcal{F} \subset \mathcal{R}^{n\times c}$ denote all the matrices with nonnegative entries. A matrix $F \in \mathcal{F}$ is a matrix that labels all points x_i with a label $y_i = \arg\max_{j\leq c} F_{ij}$. Define the series $F(t + 1) = \alpha SF(t) + (1-\alpha)\, Y$ with $F(0) = Y, \alpha \in (0, 1)$. The entire algorithm is defined as follows:

1. Form the affinity matrix $W_{ij} = exp(-\|x_i - x_j\|^2/(2\sigma^2))$ if $i \neq j$ and 0 otherwise. σ determines how fast the distance function decays.

2. Compute $S = D^{-1/2}\, W\, D^{-1/2}$ with $D_{ii} = \sum_{j=1}^{n} W_{ij}$ and $D_{ij} = 0$, $i \neq j$.

3. Compute the limit of series $\lim_{t\to\infty} F(t) = F^* = (I - \alpha S)^{-1} Y$. $\alpha \in (0,1)$ limits how much the information spreads from one point to the other.

4. Label each point x_i as $\arg\max_{j \leq c} F^*_{ij}$.

The regularization framework for this method follows. The cost function associated with the matrix F with regularization parameter $\mu > 0$ is defined as

$$\mathcal{Q}(F) = \frac{1}{2}\left(\sum_{i,j=1}^{n} W_{ij}\left\|\frac{1}{\sqrt{D_{ii}}}F_i - \frac{1}{\sqrt{D_{jj}}}F_j\right\|^2 + \mu\sum_{i=1}^{n}\|F_i - Y_i\|^2\right) \tag{1}$$

The first term is the smoothness constraint that associates a cost with change between nearby points. The second term, weighted by μ, is the fitting constraint that associates a cost for change from the initial assignments. The classifying function is defined as $F^* = \arg\min_{F\in\mathcal{F}} \mathcal{Q}(F)$. Differentiating $\mathcal{Q}(F)$ one obtains $F^* - \frac{1}{1+\mu}SF^* - \frac{\mu}{1+\mu}Y$. Define $\alpha = \frac{1}{1+\mu}$ and $\beta = \frac{\mu}{1+\mu}$ (note that $\alpha + \beta = 1$ and the matrix $(I-\alpha S)$ is non-singular) one can obtain

$$F^* = \beta(I - \alpha S)^{-1} Y \tag{2}$$

For a more in-depth discussion about the regularization framework and how to obtain the closed form expression F^* see Zhou et al. (2004).

An unlabeled point is assigned to the class with the highest value in its row of F^*, a $n \times c$ matrix. Note that the label assignment for each point depends on the initial marked points chosen and the parameters σ and α. In most cases one of the columns in F^* is significantly larger than any other value for this point indicating a clear vote for one class. Since this depends on the parameters chosen and it is not obvious how to choose the parameters, we obtained several sets of labels by varying σ which defines the local neighborhood of a point.

References

Aalsma MC, Lapsley D (2001) A typology of adolescent delinquency: sex differences and implications for treatment. Crim Behav Ment Health 11:173–191

Aldenderfer MS, Blashfield RK (1984) Cluster analysis, quantitative applications in the social sciences, vol 44. Sage, Beverly Hills, CA

Alterman AI, McDermott PA, Cacciola JS, Rutherford MJ, Boardman CR (1998) A typology of antisociality in methadone patients. J Abnorm Psychol 107(3):412–422
American Psychiatric Association (1994) Diagnostic and statistical manual of mental disorders, 4th edn. Author, Washington, DC
Arabie P, Hubert LJ, DeSoete G (eds) (1996) Clustering and classification. World Scientific, River Edge, NJ
Bergman LR (1988) You can't classify all the people all the time. Multivariate Behav Res 23:425–441
Beutler LE, Malik ML (2002) Rethinking the DSM: a psychological perspective. American Psychological Association, Washington, DC
Blackburn R (1986) Patterns of personality deviation among violent offenders—replication and extension of an empirical taxonomy. Brit J Crim 26:254–269
Breiman L (1996) Bagging predictors. Mach Learn 24:123–140
Brennan T (1987) Classification: an overview of selected methodological issues. In: Gottfredson DM, Tonry M (eds) Prediction and classification: criminal justice decision making. University of Chicago Press, Chicago, pp 201–248
Brennan T, Dieterich W (2003) Youth COMPAS psychometrics: reliability and validity. Northpointe Institute for Public Management, Traverse City, MI
Brewer WF (1993) What are concepts? Issues of representation and ontology. In: Nakamura G, Medin D, Taraban R (eds) Categorization by humans and machines. Academic Press, San Diego, CA, pp 495–533
Britt CL (1994) Versatility. In: Hirschi T, Gottfredson MR (eds) The generality of deviance. Transaction, New Brunswick, NJ, pp 173–192
Bronfenbrenner U (1979) The Ecology of human development: experiments by nature and design. Harvard University Press, Cambridge Mass
Bryant R (2000) Discovery and decision: exploring the metaphysics and epistemology of scientific classification. Fairleigh Dickinson University Press, Madison, NJ
Calinski T, Harabasz J (1974) A dendrite method for cluster analysis. Comm Stat 3:1–27
Costa PT, Herbst J, McCrae R, Samuels J, Ozer DJ (2002) The replicability and utility of three personality types. Eur J Pers 16:73–87
Duda R, Hart P (1973) Pattern classification and scene analysis. John Wiley and Sons, New York
Farrington DP (2003) Developmental and life course criminology: key theoretical and empirical issues. Criminology 41:221–255
Frick P (2004) Developmental pathways to conduct disorder: implications for serving youth who show severe aggressive and antisocial behavior. Psychol Schools 41:823–834
Gibbons DC (1975) Offender typologies—2 decades later. Brit J Crim 15:140–156
Gordon AD (1999) Classification. Chapman and Hall, NewYork
Gottfredson MR, Hirschi T (1990) A general theory of crime. Stanford University Press, Stanford, CA
Han J, Kamber M (2000) Data mining: concepts and techniques. Morgan Kaufmann, New York
Hare RD (1996) Psychopathy: a clinical construct whose time has come. Crim Justice Behav 23:25–54
Harris P, Jones PR (1999) Differentiating delinquent youths for program planning and evaluation. Crim Justice Behav 26:403–434
Hastie T, Tibshirani R, Narasimhan B, Chu G (2006) Pamr: Pam: prediction analysis for microarrays. R package version 1.29
Hirschi T, Gottfredson MR (1994) The generality of deviance. In: Hirschi T, Gottfredson MR (eds) The generality of deviance. Transaction Publishers, pp 1–22
Huizinga D, Esbensen F, Weiher AW (1991) Are there multiple paths to delinquency? J Crim Law Criminol 82:83–118
Jefferson TW, Johnson JH (1991) Relationship of hyperactivity and sensation seeking to delinquent subtypes. Crim Justice Behav 18:195–201
Jesness CF (1988) The Jesness inventory classification system. Crim Justice Behav 15:78–91
Jessor R, Donovan JE, Costa FM (1991) Beyond adolescence: problem behavior and young adult development. Cambridge University Press, New York
Jones PR, Harris PW (1999) Developing an empirically based typology of delinquent youths. J Quant Criminol 15:251–276
Lenzenweger M (2004) Consideration of the challenges, complications, and pitfalls of taxometric analysis. J Abnorm Psychol 113:10–23
Loeber R (1996) Developmental continuity, change, and pathways in male juvenile problem behaviors and delinquency. In: Hawkins JD (ed) Delinquency and crime: current theories. Cambridge University Press, New York, pp 1–28
Loeber R, Keenan K, Zhang Q (1997) Boys' experimentation and persistence in developmental pathways toward serious delinquency. J Child Fam Stud 6:321–357

Lykken D (1995) The antisocial personalities. Lawrence Erlbaum, Hillsdale NJ
McIntyre RM, Blashfield RK (1980) A nearest-centroid technique for evaluating the minimum-variance clustering procedure. Multivariate Behav Res 15(2):225–238
Mealey L (1995) The sociobiology of sociopathy: an integrated evolutionary model. Behav Brain Sci 18:523–599
Meehl PE, Yonce LJ (1994) Taxometric analysis: I. Detecting taxonicity with two quantitative indicators using means above and below a sliding cut (MAMBAC procedure). Psychol Rep 74(3):1059–1274
Megargee E, Bohn M (1979) Classifying criminal offenders: a new system based on the MMPI. Sage Beverly Hills, CA
Mezzich AC, Coffman G, Mezzich JE (1991) A typology of violent delinquent adolescents. J Psychiatry Law 19:63–78
Miller W (1958) Lower-class culture as a generating milieu of gang delinquency. J Soc Issues 14:5–19
Miller M, Kaloupek DG, Dillon AL, Keane TM (2004) Externalizing and internalizing subtypes of combat-related PTSD: a replication and extension using the PSY-5 scales. J Abnorm Psychol 113:636–645
Milligan GW (1996) Clustering validation: results and implications for applied analyses. In: Arabie P, Hubert L, DeSoete G (eds) Lustering and classification. World Scientific Press, River Edge, NJ, pp 345–379
Milligan GW, Cooper MC (1985) An examination of procedures for determining the number of clusters in a data set. Psychometrika 50:159–79
Moffitt TE (1993) Adolescence-limited and life-course persistent antisocial behavior: a developmental taxonomy. Psychol Rev 100:674–701
Moffitt TE (2003) Life-course-persistent and adolescent-limited antisocial behavior: a 10-year research review and a research agenda. In: Lahey BB, Moffitt TE, Caspi A (eds) Causes of conduct disorder and juvenile delinquency. Guilford Press, Thousand Oaks, CA, pp 49–75
Moffitt TE, Caspi A, Rutter M, Silva PA (2001) Sex differences in antisocial behaviour. Cambridge University Press, Cambridge Mass
Nagin DS, Land KC (1993) Age, criminal careers, and population heterogeneity: specification and estimation of a nonparametric, mixed Poisson model. Criminology 31:327–362
Nagin D, Paternoster R (2000) Population heterogeneity and state dependence: state of the evidence and directions for future research. J Quant Criminol 16:117–144
Newman J, MacCoon DG, Vaughn LJ, Sadeh N (2005) Validating a distinction between primary and secondary psychopathy with measures of Gray's BIS and BAS constructs. J Abnorm Psychol 114: 319–323
Osgood DW (2005) Making sense of crime and the life-course. Ann Am Acad Polit Soc Sci 602:196–211
Osgood DW, Johnston LD, O'Malley PM, Bachman JG (1988) The generality of deviance in late adolescence and early adulthood. Am Sociol Rev 53:81–93
Paternoster R, Brame R (1997) Multiple routes to delinquency? A test of developmental and general theories of crime. Criminology 35:49–84
Potter CC, Jenson JM (2003) Cluster profiles of multiple problem youth: mental health problem symptoms, substance use, and delinquent conduct. Crim Justice Behav 30:230–250
Quay HC (1990) Patterns of delinquent behavior. In: Quay HC (ed) Handbook of juvenile delinquency. Wiley-Interscience
R Development Core Team (2006) R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria
Raftery A, Dean N (2004) Variable selection for model-based clustering, Technical Report 452, Department of Statistics, University of Washington
Raine A, Moffitt TE, Caspi A (2005) Neurocognitive impairments in boys on the life-course persistent antisocial path. J Abnorm Psychol 114:38–49
Rosch E (1978) Principles of categorization. In: Rosch E, Lloyd BB (eds) Cognition and categorization. Erlbaum, Hillsdale, NJ, pp 27–48
Sampson RJ, Laub JH (2005) A life-course view of the development of crime. Ann Am Acad Polit Soc Sci 602:12–45
Simourd DJ, Hoge RD, Andrews DA, Leschied AW (1994) An empirically-based typology of male young offenders. Can J Criminol 36:447–461
Skilling TA, Quinsey VL, Craig WM (2001) Evidence of a taxon underlying serious antisocial behavior in boys. Crim Justice Behav 28:450–470
Skinner HA (1981) Toward the integration of classification theory and methods. J Abnorm Psychol 90: 68–87
Sorensen E, Johnson E (1996) Subtypes of incarcerated delinquents constructed via cluster analysis. J Child Psychol Psychiatry 37:293–303

Stefurak T, Calhoun GB, Glaser BA (2004) Personality typologies of male juvenile offenders using a cluster analysis of the Millon adolescent clinical inventory introduction. Int J Offender Ther Comp Criminol 48:96–110

Thagard P (1992) Conceptual revolutions. Princeton University Press, Princeton, NJ

VanVoorhis P (1994) Psychological classification of the adult male prison inmate. State University of New York Press, Albany, NY

VanVoorhis P (1988) A cross classification of 5 offender typologies—issues of construct and predictive-validity. Crim Justice Behav 15:109–124

Walsh A (2002) Biosocial criminology. Anderson Press, Cincinnati

Warren MQ (1971) Classification of offenders as an aid to efficient management and effective treatment. J Crim Law Criminol Police Sci 62:239–258

Wishart D (2003) Clustan graphics primer—a guide to cluster analysis, 3rd edn. Clustan Limited, Edinburgh, Scotland

Zhang L, Welte JW, Wieczorek WF (2002) Underlying common factors of adolescent problem behaviors. Crim Justice Behav 29:161–182

Zhou D, Bousquet O, Lal TN, Weston J, Schölkopf B (2004) Learning with local and global consistency. In: Thrun S, Saul L, Schölkopf B (eds) Advances in neural information processing systems 16, MIT Press, Cambridge Mass