9-17-2010

Tracing Prospective Profiles of Juvenile Delinquency and Non-Delinquency: An Optimal Classification Tree Analysis

Fred B. Bryant
Loyola University Chicago, fbryant@luc.edu

Hideo Suzuki
Loyola University Chicago

John D. Edwards
Loyola University Chicago

Recommended Citation

Bryant, Fred B.; Suzuki, Hideo; and Edwards, John D.. Tracing Prospective Profiles of Juvenile Delinquency and Non-Delinquency: An Optimal Classification Tree Analysis. Optimal Data Analysis, 1, : 125-143, 2010. Retrieved from Loyola eCommons, Psychology: Faculty Publications and Other Works,

Optimal Data Analysis
2010, Vol. 1, Release 1 (September 17, 2010)

2155-0182/10/$3.00

Tracing Prospective Profiles of Juvenile Delinquency and Non-Delinquency: An Optimal Classification Tree Analysis

Hideo Suzuki, Ph.D., Fred B. Bryant, Ph.D., and John D. Edwards, Ph.D.
Loyola University Chicago

This study explored multiple variables that influence the development of juvenile delinquency. Two datasets of the National Youth Survey, a longitudinal study of delinquency and drug use among youths from 1976 and 1978, were used: 166 predictors were selected from the 1976 dataset, and later self-reported delinquency was selected from the 1978 dataset. Optimal data analysis was then used to construct a hierarchical classification tree model tracing the causal roots of juvenile delinquency and non-delinquency. Five attributes entered the final model and provided 70.37% overall classification accuracy: prior self-reported delinquency, exposure to peer delinquency, exposure to peer alcohol use, attitudes toward marijuana use, and grade level in school. Prior self-reported delinquency was the strongest predictor of later juvenile delinquency. These results highlight seven distinct profiles of juvenile delinquency and non-delinquency: lay delinquency, unexposed chronic delinquency, exposed chronic delinquency, unexposed non-delinquency, exposed non-delinquency, unexposed reformation, and exposed reformation.

The Federal Bureau of Investigation (FBI) reported that more than 1.5 million juveniles under the age of 18 were arrested in 2003, suggesting that about 16.3% of all individuals arrested were juveniles.[1] As a result, youth violent crime is often considered to be a major problem in the United States.[2] In addition, research indicates that a delinquent criminal career increases the potential to commit crime in adulthood.[3-11] For these reasons, juvenile delinquency and its causes have been major topics in the study of crime.[12]

Some scholars have focused on situational factors as underlying determinants of criminal behavior.[13-16] For example, because crime rates are generally high in areas of poverty, it has been argued that poor socialization (i.e., failure to teach skills to achieve middle-class success) provided by lower-class parents is a predictor of delinquency.[17] With poor socialization, lower-class adolescents feel frustrated and develop a unique subculture for their values.

2155-0182/10/$3.00

From the general view of conventional groups, this is referred to as a delinquent subculture, and youths belonging to this subculture are socially labeled as delinquent gangs. Moreover, a delinquent subculture often develops in socially disorganized areas.[18] Social disorganization is said to exist[12] when: "institutions of social control... have broken down and can no longer carry out their expected or stated functions" (p. 168). Adolescents living in socially disorganized areas have limited conventional opportunities, such as well-paying jobs or educational opportunities, which adolescents eventually perceive as an unequal distribution of power, a disjunction existing between aspirations and expectations, or a discrepancy between expectations and achievement.[18] To achieve their goals under such limited conventional opportunities, some adolescents seek alternative but illegal ways and thereby become involved in a deviant subculture.

Although prior research[17-18] addressed the general relationship between social class and delinquency, not all lower-class youths automatically engage in illegal behaviors. As an alternative conceptual viewpoint, social learning theory argues that crime results from the learning process of rewarded and punished behaviors shaped through past experience and observations.[19-21] For instance, youth might learn actual criminal techniques (e.g., how to steal things from others), psychological coping strategies (e.g., how to deal with guilt or shame as a result of criminal activities), and attitudes about crime (e.g., the norms and values related to criminal activities) from direct exposure to antisocial behavior[22-23] or from relationships with a delinquent group.[24-27]

Furthermore, it has been suggested that criminals are at lower stages of moral development than law-abiding citizens.[28-30] This reasoning suggests that people's perceptions of their environment influence moral development. In fact, Thornberry[26] found that peer influence was a crucial element during mid-adolescence, and having delinquent peers helped form delinquent values. Menard and Elliott[31] also found that antisocial behavior attenuated a sense of social morality.

Considering influences that move youth away from antisocial behavior, in contrast, Hirschi[32] focused on four important prosocial bonds that detach adolescents from delinquency: attachment (i.e., sensitivity to and interest in others); involvement (e.g., participation in social activities); commitment (i.e., investing time, energy, and effort in conventional behaviors); and belief (i.e., respecting social values). According to his social bond theory, if youths have weak bonds of attachment, involvement, commitment, and belief, then they are more likely to engage in delinquent behavior. Extending this theoretical model, social bond theory was transformed into the general theory of crime (GTC), in which impulsive adolescents who receive poor socialization are more likely to be low in self-control and to weaken their social bonds to conventional groups, which, in turn, encourages them to seek criminal opportunity (e.g., joining gangs, using illegal drugs).[33]

Contrary to theoretical predictions, however, it has been reported that some youths who did not actually reject social bonds nevertheless developed associations with delinquents.[24] Thus, it is suggested that a relationship between social bonds and delinquent behavior is moderated by other factors, such as socioeconomic status.[24] Alternatively, path analyses of the National Youth Survey from 1976 to 1978 concluded that prior delinquency and involvement in delinquent peer groups were direct causal influences on delinquency and drug use, and conventional bonds and strain in-directly influenced later delinquency.[24] This research implies that delinquency is recidivistic probably because such youth have been labeled negatively and stigmatized, making it difficult for them to be rehabilitated into conventional society.[34-35]

Thus, previous research has provided rich information explaining sociological and

2155-0182/10/$3.00

psychological mechanisms underlying delinquency. Our goal in this study is to combine previous theoretical perspectives and research findings to examine delinquency more comprehensively than has been done previously. Most prior research has examined only bivariate or linear relationships with delinquency and has analyzed a limited number of predictors. In this study, we investigated many different potential predictors in a single integrated model and explored how these various predictors interact non-linearly with each other. We hypothesized that both social and personal factors would mutually influence delinquent behaviors. We also considered several personal, social, and family-related variables that are potentially associated with delinquency, such as attitudes toward deviance, social isolation, family isolation, and demographic characteristics. Our dependent variable was youth's delinquency status—delinquency versus non-delinquency—and we used a newly available non-linear multivariable method of classification tree analysis, based on optimal data analysis (ODA), to classify observations into delinquents or nondelinquents.[36]

Advantages of Classification Tree Analysis (CTA)

Traditionally, linear classification methods such as discriminant analysis and logistic regression analysis have been used to solve statistical classification problems. Nevertheless, linear classification methods have several weak points that might produce statistical solutions that are less than optimal. For example, discriminant analysis can produce probabilities beyond the range of 0 to 1 and requires restrictive normality on the independent variables, which is usually not met in practice.[37] Furthermore, both discriminant analysis and logistic regression analysis simplify complex real-world phenomena by using a linear model although real phenomena are typically not linear.[38] In addition, these linear methods assume three conditions that are often unrealistic—namely, that the magnitude of importance, the direction of influence, and the coefficient value for each predictor variable is the same across all observations.[38] *It is not our intention to argue that statistical results found by linear methods are invalid, but rather to note that the level of accuracy of these methods is constrained by the above limitations.*

In contrast to traditional linear classification techniques, the ODA paradigm offers a non-linear multivariable classification method known as hierarchically optimal classification tree analysis (CTA).[38] Independent and dependent variables are referred to respectively as "attributes" and "classes" in CTA. An attribute is defined as: "any variable that can attain two or more levels, and reflects the phenomenon that one hopes will successfully predict the class variable," and a class variable is defined as "any variable that can attain two or more levels, and reflects the phenomenon that one desires to successfully predict."[36]

Note that a class variable must be categorical, although an attribute can be either categorical or continuous. CTA has distinct advantages over linear classification methods. First, CTA can handle non-linear, complicated real-world phenomena. With CTA, the shape or form of a given phenomenon does not matter, whereas linear methods assume that a straight line or a sigmoidal curve characterizes the underlying phenomenon.[38] In addition, a CTA model produces a high level of classification accuracy by adopting optimal decision rules, rather than trying to maximize explained variance or minimize a fit function (see Method for more detail). Moreover, CTA is free from the restrictive assumptions about independent variables. In particular, unlike linear methods, CTA does not assume constant importance, direction of influence, and coefficient value (unstandardized or standardized regression coefficient) for each attribute across all observations.[38]

Another strength of CTA is it provides a hierarchically optimal classification model, which can be very informative. In CTA, the at-

2155-0182/10/$3.00

tribute with the strongest effect size for the total sample, called the first node, enters the top of a hierarchically optimal classification tree model. One level or branch of the first node leads to a second node through a predictive pathway, while another level of the first node leads to another second node through a different predictive pathway. At these second nodes, the attributes with the strongest effect size under each condition are entered to produce, in turn, different pathways to the third nodes. These patterns are repeated until prediction endpoints are reached.

The final CTA model reveals two important pieces of information. First, tracing combinations of nodes in CTA visually identifies crucial interaction effects. For example, imagine the final CTA model indicates a certain subgroup (endpoint) is predicted to engage in delinquency when the first node of the model (e.g., attachment) is at a low value and the second node (e.g., moral belief) is also low. This result indicates that moral belief predicts delinquency, depending on the strength of attachment. Note that in contrast to traditional linear approaches, CTA automatically detects important interactions by examining all attributes in the statistical model. Second, the CTA model allows us to trace multiple stages branching into each level of a class variable and to discover the critical profiles linked to each outcome. In the above example, the CTA model would show attachment (the first stage) and moral belief (the second stage) at which youths move toward delinquency or non-delinquency. This result implies that one profile of delinquency is the combination of weak attachment and moral beliefs.

In contrast, linear methods cannot identify ordinal predictors leading to each outcome. Furthermore, unlike CTA, linear methods have difficulty finding combinations of multiple variables predicting each level of an outcome simultaneously, making it more difficult to use linear methods to identify predictive profiles.

These advantages make CTA a powerful procedure for solving statistical classification problems in comparison with the linear classification methods. CTA models are manually constructed using statistical software which conducts ODA and classifies observations optimally by following "a prediction rule that explicitly achieves the theoretical maximum possible level of classification accuracy".[36] We used ODA in this study for three reasons in addition to the fact that ODA enables us to capitalize on all the strengths of CTA. First, ODA can analyze all types of attributes measured by ratio, interval, ordinal, and nominal scales.[36,39] Second, as noted in the Method section below, ODA empirically tests the expected cross-sample generalizability of optimal classification models.[36,39] Finally, ODA simultaneously analyzes as many attributes as one wants without the limitations of the ratio of attributes to sample size or problems of multicollinearity.[36] This is because ODA tests the overall effect of each attribute on a class variable individually and selects only the single most influential attribute at each node. This strategy differs from multiple regression analysis, which calculates the partial effect of each variable independent of the effects of other variables when considered simultaneously.

Method

Participants and Materials. Archival data from the National Youth Survey, a 1976-1978 longitudinal design with multiple birth cohorts, were used.[24,40-41] In early 1977, the first wave of the survey gathered a multistage, cluster (area) probability sample of 1,725 American adolescents aged from 11 to 17 in 1976. Thus, by design, the sample included not only delinquents but also non-delinquents. The survey assessed events and behaviors theoretically linked with delinquency during calendar year 1976, and the subsequent wave tracked most of the individuals in 1978. Because the National Youth Survey followed the same individuals over time, we selected theoretically relevant attributes from the 1976 dataset to predict later self-reported delinquency in the 1978 dataset. Partici-

2155-0182/10/$3.00

pants interviewed for the first survey were representative of the youth population aged 11-17 in the U.S. measured by the U.S. Census Bureau, and the attrition rate for the subsequent wave was only 6% (N=99).[24] ODA software[36] was used to manually construct a hierarchically optimal CTA model of juvenile delinquency.

Measures. Our class variable of general delinquency was a composite index consisting of the frequency of the following behaviors reported by youths in 1978: aggravated assault, larceny, burglary, robbery, marijuana use, hallucinogens use, amphetamines use, barbiturates use, cocaine use, vandalism, buying stolen goods, hitting, joyriding, runaway, carrying a hidden weapon, prostitution, and selling drugs. Note that there were no questions about homicide and arson in the survey. Alcohol use, lying about age, hitchhiking, and buying liquor for a minor from were excluded from our measure of delinquency because they were rather common illegal acts.[24,43] Sexual intercourse, panhandling, and disorderly conduct were also excluded from delinquent behaviors. Sexual intercourse is relatively commonplace among youths, and it is also hard to judge whether sexual intercourse is delinquent.[43] For example, a victim of rape has sexual intercourse against his or her will, but voluntary intercourse is not illegal. Thus, it was reasonable to bar sexual intercourse as a component of delinquency. As for panhandling, begging for money does not hurt anyone and is not delinquent. Finally, people often behave in a disorderly manner (e.g., being loud in public) simply because of their exuberantly positive mood, so disorderly conduct is not always a form of delinquency.

Although our decision to consider some illegal acts as non-delinquent due to the trivial nature of these acts may not be universally accepted, the proportion of youths who performed at least one of these "trivial" illegal acts once or more monthly was 69.1%, whereas the proportion of youths who committed delinquent acts once a month or more as we have operationally defined this construct was 32.8%, which seems much more reasonable as an estimate of the underlying rate of delinquency.

The National Youth Survey offered two sets of questions to measure (a) the actual number of times each delinquent act was committed and (b) the frequency of each delinquent behavior using a scale ranging from one (never) to nine (two-three times a day). Cronbach's α for the frequency rates of the general delinquency was 0.713, which was greater than that for the actual number of delinquent behaviors. Hence, only the frequency rate items were used to construct the class variable for CTA. Committing each delinquent behavior once a month or more (score≥4) was recoded as one point, while committing each delinquent behavior less than once a month (score<4) was recoded as zero points. This rule was the most effective in making our sample as representative as possible of American delinquents and non-delinquents (see the above discussion of the proportion of delinquents). Respondents who scored at least one point were defined as delinquents, whereas respondents who scored zero points were defined as non-delinquents: this was the class variable employed in CTA.

Attributes. A total of 166 attributes were examined, including 17 theoretical "broad band" composite variables, the individual "narrow band" items composing these theoretical attributes, and additional background and demographic characteristics used in prior research.[24] The theoretical variables were: (a) *conventional involvement* measured by a sum of scores on the school athletic and activities involvement scales and community involvement scale (α=0.70); (b) *attachment to family* measured by a sum of scores on the family involvement and aspiration scales (α=0.72); (c) *conventional commitment* measured by a sum of scores on the school aspirations scale and future occupational and educational goal scales (α=0.71); (d) *moral belief* measured by a sum of scores on the family, school, and peer normlessness scales (α=0.72);

2155-0182/10/$3.00

(e) *exposure to peer delinquency* measured by a sum of scores on the number of close friends performing each of some bad behaviors (α=0.82); (f) *involvement with delinquent peers* measured by a sum of scores on the peer involvement scale multiplied by the difference between an observed score for exposure to peer delinquency and its mean (because this is a single index, α was not computed[24]); (g) *socialization* measured by a sum of scores on the perceived sanctions in family scale (α=0.84); (h) *attitudes toward deviance* measured by a sum of scores on the attitudes toward deviance scale (α=0.79); (i) *social disorganization* measured by a sum of scores on the neighborhood problems scale and the reversed and standardized family income scale (α=0.75); (j) *prior self-reported delinquency* measured by a sum of scores on the continuous frequency rate scale (α=0.95) and measured by a sum of scores on the dichotomous frequency rate scale (α=0.91); (k) *social isolation* measured by a sum of scores on the family and school social isolation scales (α=0.73); (l) *family isolation* measured by a sum of scores on the family social isolation scale (α=0.72); (m) *social labeling* measured by a sum of scores on the family and school labeling scales (α=0.86); (n) *perceived labeling by parents* measured by a sum of scores on the family labeling scale (α=0.71); (o) *perceived labeling by teachers* measured by a sum of scores on the school labeling scale (α=0.80); and (p) *strain* measured by a sum of scores recoded 0 (no strain) to 3 (high level of strain), after subtracting scores on the achievement of each goal from scores on the importance of the corresponding goal (α=0.62).[24] Note that in measuring prior delinquency based on both continuous and dichotomous scales, we adopted the same operational definition as that of our class variable.

Procedure and Analysis Strategy. The National Youth Survey data sets were obtained through the Inter-University Consortium for Political and Social Research (ICPSR) of the University of Michigan. After all data were accessed and gathered, the class variable and attributes were selected and computed as described above. Finally, the class variable and the attributes were input into the ODA program to construct the CTA model.

To facilitate clarity of exposition we review how optimal data analysis operates in constructing a CTA model. ODA is first used to determine a cutpoint, or decision rule, for each attribute that maximizes the overall percentage of observations that are correctly classified (i.e., the percentage accuracy in classification, or PAC). For each equal interval or ordinal (i.e., continuous) predictor, ODA identifies an optimal classification cut-point (e.g., if age>14, then predict delinquency; if age≤14, then predict non-delinquency) that maximizes overall PAC. For each nominal or binary (i.e., categorical) predictor, ODA identifies an optimal classification rule (e.g., if ethnicity=Anglo, then predict delinquency; if ethnicity≠Anglo, then predict non-delinquency) that maximizes overall PAC. Thus, ODA can accommodate multi-category nominal predictors, such as race, without dummy coding these variables. Unlike other statistical methods for constructing tree models (e.g., regression-based CART or chi-square-based CHAID), ODA uses an exact permutation probability with no distributional assumptions, assesses the expected cross-sample generalizability of classification rules through an automated jackknife validity analysis procedure, and finds main effects and nonlinear interactions that optimally classify admission decisions. PAC is computed as 100% x (number of correctly classified observations)/(total number of observations).[36]

After determining the optimal cutpoint providing the greatest PAC for each attribute, the next step is to decide which attributes to enter into the hierarchically optimal CTA model. The chosen attribute must have the greatest effect strength for sensitivity (ESS), which reflects how much better PAC is compared to chance, using a standardized scale where chance

2155-0182/10/$3.00

classification accuracy is 0% and perfect classification accuracy is 100%. ESS is calculated using the following equation:

$$\mathrm{ESS}\ (\%) = \left\{1 - \frac{100 - (\text{mean PAC across classes})}{100 - \dfrac{100}{C}}\right\} \times 100$$

where C is the number of response categories for the class variable.[36] By rule-of-thumb, ESS values < 0.25 are regarded as weak, values between 0.25 and 0.50 are considered moderate, and values > 0.50 are defined as strong.[36]

After selecting the attribute with the greatest ESS to serve as a node of a tree model, the attribute's expected cross-sample stability in classification performance is assessed using a leave-one-out (LOO), or jackknife, validity analysis. In LOO analysis, classification performance is evaluated after removing an observation, and then the removed observation is classified again according to the classification performance obtained using the remaining subsample. This process is repeated until every observation has been removed and classified. An attribute is included in the CTA model only if its classification accuracy is stable in LOO analysis. LOO analysis helps to construct a tree model whose constituent attributes are most likely to generalize to a new sample.

If a LOO stable attribute with the greatest ESS is statistically significant, then the attribute enters as the first node of a CTA model. The level of statistical significance is determined by Monte Carlo simulation as a permutation probability, and is isomorphic with Fisher's exact *p* test for binary attributes. After the first node is determined, ODA subsequently searches the second node and lower nodes under each level of the highest node of a hierarchical tree model using the above procedures. These procedures are repeated until no more attributes are below the critical $p<0.05$-level.

Note that a given attribute can re-enter a node at a lower level even if it has already entered as a node at a higher level in the CTA model. This is the case when a re-entered attribute still contributes to the best classification performance with a new cutpoint when combining specific levels of higher nodes. Finally, to control the experimentwise Type I error rate at $p<0.05$ per comparison, a sequentially-rejective Sidak Bonferroni-type multiple comparisons procedure is used to prune attributes selected by inflation of Type I error.[36] These adjustments also help maximize statistical power by rejecting lower nodes tested from very small subsample sizes when the total sample becomes divided and reduced.[36]

Results

Univariate Analyses. To describe simple relationships between delinquency and each attribute, we first conducted univariate analyses using ODA (Table 1). Consistent with previous findings, most theoretical attributes were significantly related to delinquency in the predicted direction: delinquency was significantly associated with weak attachment to family, weak conventional commitment, weak moral belief, greater exposure to peer's delinquency, positive attitudes toward deviance, high level of social disorganization, more experiences of prior delinquency, high level of social isolation, high level of family isolation, negative social labeling, negative social labeling by teachers, and high level of strain.

In addition to these theoretical attributes, race and age were also significantly related to delinquency: Anglo adolescents were more likely to commit delinquency than other racial groups; and adolescents aged 14 or older were more likely to commit delinquency than those aged 13 or younger.

2155-0182/10/$3.00

Table 1: Univariate Associations of Theoretical and Demographic Attributes with Delinquent (1) Versus Non-Delinquent Behavior (0) for the Total Sample (N=1,606)

Attribute	ODA Model	*n*	% Delinquent	ESS	*p*-value
Conventional involvement	> 20.5, predict 0	70	30.00	17.93	0.413
	≤ 20.5, predict 1	186	36.56		
Attachment with family	> 29.5, predict 0	1024	25.78	19.94	0.118×10^{-13}
	≤ 29.5, predict 1	536	45.15		
Conventional commitment	> 30.0, predict 0	875	24.00	21.38	0.906×10^{-15}
	≤ 30.0, predict 1	705	42.98		
Moral belief	> 42.5, predict 0	907	25.58	18.95	0.935×10^{-12}
	≤ 42.5, predict 1	653	42.73		
Exposure to peer's delinquency	≤ 16.5, predict 0	809	21.88	30.96	0.102×10^{-26}
	> 16.5, predict 1	538	50.56		
Involvement with delinquent peers	≤ 1.26, predict 0	812	21.80	31.19	0.107×10^{-25}
	> 1.26, predict 1	532	50.75		
Socialization	> 30.5, predict 0	57	26.32	1.08	0.175
	≤ 30.5, predict 1	1520	33.16		
Attitudes toward deviance	> 25.5, predict 0	878	21.75	27.32	0.524×10^{-24}
	≤ 25.5, predict 1	719	46.04		
Social disorganization	≤ 12.15, predict 0	1377	31.30	3.79	0.0112
	> 12.15, predict 1	135	41.48		
Prior self-reported delinquency	≤ 33.5, predict 0	1053	20.42	36.86	0.215×10^{-46}
	> 33.5, predict 1	553	56.42		

2155-0182/10/$3.00

Social isolation	≤ 20.5, predict 0	662	29.15	6.49	0.0082
	> 20.5, predict 1	917	35.01		
Family isolation	≤ 10.5, predict 0	1018	29.76	8.59	0.000519
	> 10.5, predict 1	577	37.95		
Social labeling	> 81.5, predict 0	1050	26.67	19.20	0.462×10^{-13}
	≤ 81.5, predict 1	479	46.35		
Perceived labeling by parents	> 37.5, predict 0	1146	28.88	13.34	0.682
	≤ 37.5, predict 1	403	44.17		
Perceived labeling by teachers	> 43.5, predict 0	1010	25.94	19.97	0.132×10^{-13}
	≤ 43.5, predict 1	541	45.29		
Strain	≤ 11.5, predict 0	171	23.98	3.66	0.0479
	> 11.5, predict 1	1095	30.50		
Exposure to peer's alcohol use	≤ 2.5, predict 0	880	22.05	32.13	0.332×10^{-30}
	> 2.5, predict 1	501	52.89		
Attitudes toward marijuana use	> 3.5, predict 0	1042	23.61	27.01	0.553×10^{-25}
	≤ 3.5, predict 1	556	49.82		
Sex	Male, predict 0	849	40.64	-18.75	0.999
	Female, predict 1	757	24.04		
Race	Black/Chicano/American Indian/Asian/other, predict 0	322	25.47	6.69	0.000902
	Anglo, predict 1	1281	34.66		
Age	≤ 13, predict 0	732	24.45	17.28	0.346×10^{-10}
	> 13, predict 1	874	39.82		

2155-0182/10/$3.00

Grade at School	8th grade or lower, predict 0	819	26.01	17.28	0.439
	9th grade or higher, not in school, or other, predict 1	787	39.90		
GPA	F, predict 0	10	60.00	-0.78	0.983
	A, B, C, or D, predict 1	1585	32.49		
Family Income	≤ $14,000, predict 0	141	33.33	-0.43	0.646
	> $14,000, predict 1	1375	32.22		
Parent's Marital Status	Single or married, predict 0	1300	31.23	5.11	0.593
	Divorced/separate/other, predict 1	280	38.93		

Note: "ODA Model" indicates the cutpoint or decision rule by which ODA classified (non)delinquents.[36] Total sample sizes varied across attributes due to incomplete data. A sequentially-rejective Bonferroni adjustment procedure was *not* employed for univariate analyses.[36] The total number of respondents who answered the set of questions associated with conventional involvement was 256, so the response rate for this set of items was only 15.94%. ESS values indicated in red were stable in jackknife ("leave-one-out") validity analysis, and are expected to show cross-sample generalizability.

However, contrary to previous theory and research, attributes unrelated to delinquency included conventional involvement, socialization, and perceived labeling by parents. Moreover, LOO analysis concluded that a significant relationship between involvement with delinquent peers and delinquency was not cross-sample generalizable.

Classification Tree Analysis. Our primary interest was not to see simple relationships between each attribute and delinquency, but to see how multiple attributes combine to explain predictive roots and profiles of juvenile delinquency and non-delinquency. Therefore, we used ODA to construct a hierarchically optimal CTA model. Following established procedures for constructing optimal CTA models, 68 nodes were initially identified; but after applying a sequentially-rejective Sidak Bonferroni-type multiple comparisons procedure, only five nodes were retained. These five nodes were prior self-reported delinquency measured by continuous scales as the first node (p<0.001) and as the third node (p<0.001), exposure to peer alcohol use during 1976 (p<0.001), exposure to peer delinquency during 1976 (p<0.001), grade level in school during 1976 (p<0.001), and attitudes toward marijuana use during 1976 (p<0.001). Except for grade level, all attributes were significant in the univariate analyses. Figure 1 shows the final hierarchically optimal CTA model for explaining juvenile delinquency. In the figure, circles represent nodes, arrows indicate branches, and rectangles are prediction endpoints (D=delinquency, ND=non-delinquency). Numbers below each node indicate directional Fisher's exact p value for the node, and numbers in parentheses within each node indicate ESS for the node. Also, numbers next to each arrow indicate the value of the cutpoint for the node.

The strongest predictor of delinquency for the total sample was prior self-reported delinquency (ESS=36.86%): the first node of the CTA model. The cutpoint for this attribute was 33.5 (1.94% on the absolute scale).

Copyright 2010 by Optimal Data Analysis, LLC
2155-0182/10/$3.00

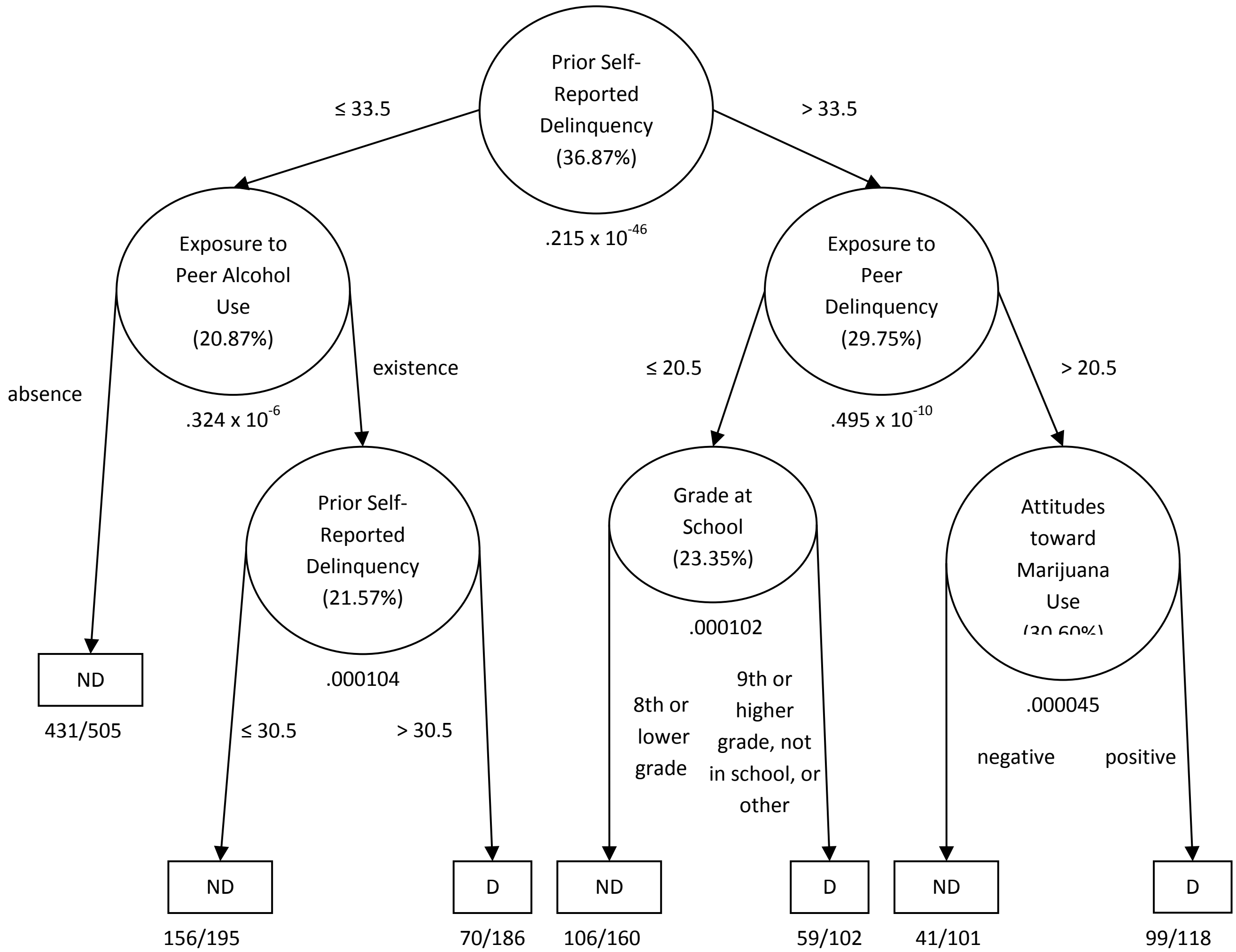

Figure 1: The CTA model for predicting juvenile delinquency versus non-delinquency (N=1,367). Ellipses represent nodes, arrows represent branches, and rectangles represent prediction endpoints. Numbers under each node indicate the exact p value for each node. Numbers in parentheses within each circle indicate effect strength. Numbers beside arrows indicate the cutpoint for classifying observations into categories (delinquency or non-delinquency) for each node. Fractions below each prediction endpoint indicate the number of correct classifications at the endpoint (numerator) and the total number of observations classified as the endpoint (denominator). Negative attitudes toward marijuana use = Thinking that marijuana use is "very wrong" or "wrong" for a youth or someone his or her age; Positive attitudes toward marijuana use = Thinking that marijuana use is "a little bit wrong" or "not wrong at all" for a youth or someone his or her age; D = delinquency; ND = non-delinquency.

2155-0182/10/$3.00

For youths who scored 33.5 or less on the prior delinquency scale based on its frequency rate, the second node was exposure to peer alcohol use (ESS= 20.87%). If a respondent had no friends who used alcohol, then that respondent was predicted to be non-delinquent with 85.35% accuracy. In other words, a few prior experiences with delinquency and no exposure to peer alcohol use jointly led to non-delinquency. For youths who had a few prior experiences of delinquency but who were exposed to peer alcohol use, a third node branched to either delinquency or non-delinquency. This third node was, again, prior self-reported delinquency (ESS=21.57%). That is, prior self-reported delinquency became the strongest attribute again among youths who had committed delinquent behavior less frequently and were exposed to peer alcohol use, but not among youths who fell into the other predictive pathways. At this node the cutpoint was 30.5, representing less than the 1st percentile on an absolute scale. If youths scored 30.5 or lower on the prior delinquency scale, then they were predicted to be non-delinquent with 80% accuracy. Therefore, even if youths had friends who had used alcohol, it was possible that the youths were still non-delinquents when they had been much less likely to perform delinquent behaviors two years earlier. In contrast, under the conditions where youths were exposed to peer alcohol use, if their scores were above 30.5 but 33.5 or less on the prior delinquency scale, then they were predicted to be delinquent with 37.63% accuracy. This was the lowest classification performance at any endpoint predicting delinquency. Overall predictive accuracy for youths who had earlier engaged in delinquent acts less often was 74.15% (657/886).

In comparison, for those who had earlier engaged in delinquent behavior more often, a different hierarchical pattern appeared. Among youths who scored more than 33.5 on the prior delinquency scale, the strongest predictor in the model was exposure to peer's delinquency. The cutpoint for this attribute was 20.5, which represents the 26th percentile on an absolute scale. If youths scored more than 20.5 on the scale of exposure to peer delinquency, then they were classified as being either delinquent or non-delinquent, depending on their attitudes toward marijuana use. In contrast, among youths reporting more frequent prior delinquency and less exposure to peer's delinquency (score≤ 20.5), classification as delinquent or nondelinquent depended on their grade level in school. Specifically, youths were predicted as non-delinquent when (a) they were more exposed to peer delinquency and thought that marijuana use was "very wrong" or "wrong" for them or someone their age (59.41% delinquency rate), or (b) they were less exposed to peer's delinquency and were in the eighth grade or lower (33.75% delinquency rate). In comparison, youths were classified into delinquency when (c) they were more exposed to peer delinquency and thought that marijuana use was "a little bit wrong" or "not wrong at all" (83.90% delinquency rate), or (d) they were less exposed to peer's delinquency and were in ninth grade or higher, did not attend at school, or a trade or business school (57.84% delinquency rate). Overall predictive accuracy for those who reported more frequent delinquent behaviors earlier was 63.41% (305/481).

Table 2 summarizes the overall classification performance of the CTA model, which correctly classified 962 (70.37%) of the total 1,367 youths. The ESS for this model was 30.59%, indicating that the model attained almost one-third of the theoretically possible improvement in classification accuracy versus the performance expected by chance: this is considered to reflect a moderate effect.[36]

2155-0182/10/$3.00

Table 2: Confusion Table for CTA DelinquencyModel

		Predicted Class Status		
		Non-Delinquent	Delinquent	
Actual Class Status	Non-Delinquent	860	128	Specificity = 87.0%
	Delinquent	135	70	Sensitivity = 34.1%
		Negative Predictive Value = 86.4%	Positive Predictive Value = 35.4%	

Additional Comments about Cutpoints. Although the cutpoints for prior self-reported delinquency were 33.5 and 30.5, depending on the level of node, what do these values signify? Scores less than 33.5 were located within 1.94% on the absolute possible range, and the scores less than or equal to 30.5 reflects 0.65% of the absolute possible range on the prior delinquency scale. Descriptive statistics showed that the mean of prior delinquency (range=29-261) was 35.02 with *SD*=15.40. Overall, 65.2% of respondents scored 33.5 or less, while 34.8% scored more than 33.5. Conceptually, a respondent who scored 29 (i.e., 1 point x 29 items) had never committed delinquency in 1976, and a respondent who had performed all types of delinquent behaviors once or twice in 1976 should have scored 58 (i.e., 2 points x 29 items). Therefore, respondents who scored 33.5 had performed only a few types of illegal behaviors once or twice in 1976. In addition, because the score of 30 indicates that a respondent committed one kind of delinquent behavior once or twice in 1976, scores less than or equal to 30.5 indicate that respondents were engaged in only one delinquent behavior very few times. Thus, scores below 33.5 on the prior delinquency index were much closer to the score of non-delinquents used to categorize the class variable, and could be considered as reporting very few prior delinquent experiences.

What about exposure to peer delinquency? The cutpoint for exposure to peer delinquency was 20.5. Descriptive statistics revealed that the mean of this attribute (range=10-50) was 16.72 with *SD*=5.87. For exposure to peer delinquency, 77.8% of respondents scored 20.5 or less, and 22.2% scored greater than 20.5. Scores less than 20.5 fell within 26.25% on an absolute scale. A score of 20 (i.e., 2 x 10 items) would indicate that a respondent was exposed to peers who committed all ten types of delinquent behaviors. Therefore, a score of 20.5 or less indicates that a respondent was exposed to relatively few delinquent peers.

Discussion

Implications of the CTA Model of Delinquency. As hypothesized, this study yielded a parsimonious model identifying social (exposure to peer alcohol use, exposure to peer delinquency, and grade level in school) and personal variables (prior delinquency and attitudes toward marijuana use) that together predicted American youths as either delinquent or non-delinquent, supporting the critical influence of these factors on young people's anti-social behavior. The optimal CTA model achieved about a third of the possible improvement in classifi-

2155-0182/10/$3.00

cation accuracy relative to chance, which represents a moderate effect size. The model identified three profiles of juvenile delinquency: (a) lay delinquency, reflecting infrequent prior delinquency with exposure to peer alcohol use (37.63% accuracy), (b) unexposed chronic delinquency, reflecting youth who had frequent prior delinquency with less exposures to peer delinquency, but being in the ninth grade or higher (57.84% accuracy), and (c) exposed chronic delinquency, reflecting youth who had frequent prior delinquency with exposure to peer delinquency and positive attitudes toward marijuana use (83.90% accuracy). In contrast, the model yielded four profiles of non-delinquency: (a) unexposed non-delinquency, reflecting youth who have infrequent prior delinquency with no exposure to peer alcohol use (85.35% accuracy), (b) exposed non-delinquency, reflecting youth who had extremely infrequent prior delinquency with exposure to peer alcohol use (80.00% accuracy), (c) unexposed reformation, reflecting youth who had frequent prior delinquency with less exposure to peer delinquency, but who were in eighth grade or lower (66.25% accuracy), and (d) exposed reformation, reflecting youth who had frequent prior delinquency with greater exposure to peer delinquency, but who had negative attitudes toward marijuana use (40.59% accuracy).

The CTA model provides additional insights into the prospective predictors of delinquency. Prior delinquency was the strongest predictor of subsequent delinquency—a conclusion that is consistent with previous reports that prior general delinquency directly influences later delinquency and drug use.[24] Our results extend prior findings, by identifying combinations of variables that exert a differential influence for experienced delinquents versus other subgroups of youth. For experienced delinquents, the factors important in maintaining delinquency appear to be exposure to peer delinquency, grade level in school, and attitude toward marijuana use. Youths who maintained their status as delinquents were categorized as unexposed or exposed chronic delinquents with 71.82% accuracy (Table 3). Previous studies showing the effect of exposure to antisocial behavior on criminal actions[22-23] and the effect of peers on the formation of delinquent values[26,31] support the profile of exposed chronic delinquency. Thus, with exposed chronic delinquency, prior delinquent experiences and exposure to delinquent peers might lead youths to form positive attitudes toward marijuana use, and these antisocial attitudes might encourage them to commit delinquent actions later. Note, however, that there is also a predictive profile reflecting exposed reformation, implying that not all youths with frequent prior delinquency and more exposure to delinquent peers automatically adopt positive attitudes toward marijuana.

In contrast, for adolescents who have infrequent prior delinquency, the variables predictive of changing non-delinquency into delinquency were exposure to peer alcohol use and prior delinquency. However, the combination of these factors predicted lay delinquency with only 37.63% accuracy, indicating that other factors not measured in the survey also operate.

Table 3: Summary of Cross-Classification by Year (N=1,367)

Year of 1978	Year of 1976: Non-Delinquency	Year of 1976: Delinquency
Non-Delinquency	587/700 (83.86%)	147/261 (56.32%)
Delinquency	70/186 (37.63%)	158/220 (71.82%)

Note. The numerator of each fraction indicates the number of observations classified correctly. The denominator of each fraction indicates the number of observations predicted as a given category by the CTA model. Percentages reflect the proportion of correctly classified observations.

2155-0182/10/$3.00

Another important implication is that the factors that maintain non-delinquency are different from the factors that terminate delinquency (Figure 1). The CTA model demonstrated that unexposed and exposed non-delinquents maintained their status of non-delinquency with 83.86% accuracy, whereas unexposed and exposed reformers became non-delinquents with only 56.32% accuracy (see Table 3). Future researchers should include measures of the variables composing these profiles, in order to enhance accuracy in predicting and understanding the dynamics of juvenile delinquency.

The CTA model identified protective factors more accurately than risk factors, and classification accuracy for non-delinquency was greater than for delinquency. This is probably because the surveys did not assess some critical risk factors. For instance, impulsivity[33], attention deficit/hyperactivity disorder[44], criminal opportunity[33.45], and historical contexts, such as a change in the level of surplus value[46] have all been identified as important risk factors, but were not directly assessed by the surveys. Another interesting implication concerns the crucial roles of adolescent exposure to peer delinquency and substance use in relation to delinquency. Regardless of prior delinquency, youths are sensitive to influence from peers perhaps because they desire to maintain intimacy and to avoid being rejected by peers. Also, alcohol use seems to be a "gateway" to performing delinquent behaviors by youths with infrequent prior delinquency, while marijuana use may be an obstacle to stopping delinquent behaviors.

Some variables found to be predictive of delinquency in previous research did not appear in the final CTA model. These predictors were socialization[17,24,33], social disorganization and social strain[18,24], involvement with delinquent peers[24-27], any types of social bonds[24,32-33], and any form of labeling.[34-35] It should be noted that in the univariate analyses all of these predictors—except for involvement with delinquent peers, conventional involvement, socialization, and perceived labeling by parents—were significantly predictive of delinquency (Table 1). The reason why these particular predictors failed to enter the final CTA model was that these predictors had smaller ESS than attributes selected for entry in the model, had low generalizability across samples, and/or had weaker effects when combined with variables in higher nodes of the hierarchical tree model. In contrast, grade in school was not significant in the univariate analysis, yet it was a node in the CTA model. This indicates that grade in school is significant among only a certain group, that is, American young people who had more prior delinquent experiences and were more likely to be exposed to peer delinquency, but not among general American young population.

Limitations. Our results are not without limitations. Although the strongest predictor of delinquency was prior self-reported delinquency, this result subsequently raises a follow-up question, "What factors, if any, predict prior delinquent behavior?" In our model, the profile of lay delinquency included not only those who had no prior delinquent experience, but also those who had very few prior delinquent experiences. Future research should explore the additional profile of delinquent youth who have no prior experiences of delinquency whatsoever.

Another limitation of the present research is the time frame of the survey data we analyzed. The National Youth Survey was conducted in 1976 and 1978. Thus, our results might reflect phenomena that are no longer generalizable to the present time period. Future research should address this limitation by constructing CTA using more recent data.

In terms of methodological limitations, our model reflects roughly 60% of the eligible youths originally selected by the multistage cluster sampling method. Although there is no agreed-upon standard for what constitutes an acceptable rate of inclusion, excluding 40% of respondents raises the possibility of potential selection and non-response biases. However, no

2155-0182/10/$3.00

particular group of the youth population appears to be over- or under-represented in our sample, compared to the original sample who agreed to participate in the National Youth Survey.[24]

Other methodological issues concern the particular measures used in the National Youth Survey. In particular, the self-report items used to assess delinquency and other socially negative behaviors might not accurately reflect the actual levels of these behaviors because of social desirability, memory limitations, and motivation to recall. Moreover, the National Youth Survey did not include some variables that we wanted to examine as potential predictors of delinquency (e.g., impulsivity). Future research needs to include measures of other unanalyzed variables so that the classification accuracy of the hierarchical tree model can be further improved. Finally, although some theoretical composite attributes showed acceptable values of Cronbach's α, other attributes, including exposure to peer alcohol use and attitude toward marijuana use, were each measured by only a single individual question and had unknown reliability. Future research should measure attributes, especially exposure to peer alcohol use and attitude toward marijuana use, using multiple items, obtain acceptable Cronbach's α for these composite subscales, and then re-test them by including them in an ODA model.

Finally, it should be noted that an alternative definition of delinquency might yield different findings concerning the prospective predictors of juvenile delinquency. Although we contend that the classification of delinquency or non-delinquency based on our definition produced representative samples of youths who engage in these two forms of behavior, other theorists or researchers might well adopt an alternative definition of these two constructs. Or, they might suggest examining more specific delinquent actions (e.g., theft) independently rather than a broader, comprehensive category of juvenile delinquency because the factors might vary across different delinquent actions. Nevertheless, while we should avoid over-generalizing the factors found in our study to all delinquent actions, it is also informative to focus on the large-scale pattern of delinquency. This macro-level analysis is important because (1) the society and citizens tend to be more interested in getting a general idea (e.g., how to prevent delinquent crime in general) than a specific idea (e.g., how to prevent each potential delinquent actions specifically), and (2) each specific delinquent action is not exclusive or independent but accompanies another illegal action (e.g., robbery and assault could occur at the same time). Thus, our findings provide an overview of delinquent behavior, and the next goal should be to focus on each specific delinquent action to examine whether our model is applicable to it.

Another limitation concerning our definition of delinquency is the inevitable loss of precision in analyzing delinquency as a dichotomy as opposed to a continuous rate of frequency. In doing so, we have limited ourselves to investigating variables that predict whether or not youths exceed a threshold frequency that we have defined a priori as representing juvenile delinquency versus non-delinquency. These predictive variables may well differ from those that explain variation in the absolute frequency of delinquent behaviors.

Applications of the Present Study. The findings suggest potentially effective strategies for crime prevention. For example, shifting positive attitudes toward marijuana use toward negative attitudes may reduce delinquent behavior among exposed but reformed delinquent youths. Furthermore, our results suggest that an effective approach to protect non-delinquent youths from moving toward delinquency is to keep them away from peers who use alcohol. Future research should test these hypotheses.

References

[1]Federal Bureau of Investigation. (2004). *Crime in the United States 2003: Uniform crime re-*

2155-0182/10/$3.00

ports. Retrieved November 9, 2004, from http://www.fbi.gov/ucr/03cius.htm

[2]Soler, M. (October, 2001). *Public opinion on youth, crime, and race: A guide for advocates*. Retrieved November 9, 2004, from http://www.buildingblocksforyouth.org/advocacyguide.html#juvcrime.

[3]Farrington, D.P. (1986). Stepping stones to adult criminal careers. In D. Olweus, J. Block, & M. Radke-Yarrow (Eds.), *Development of antisocial and prosocial behavior* (pp. 359-384). New York: Academic Press.

[4]Farrington, D.P. (1989). Early predictors of adolescent aggression and adult violence. *Violence and Victims*, *4*, 79–100.

[5]Farrington, D.P. (1991). Childhood aggression and adult violence: Early precursors and later-life outcomes. In D.J. Pepler & K.H. Rubin (Eds.), *The development and treatment of childhood aggression* (pp. 5-29). Hillsdale, NJ: Erlbaum.

[6]Farrington, D.P. (1995). The development of offending and antisocial behavior from childhood: Key findings from the Cambridge study in delinquent development. *Journal of Child Psychology and Psychiatry*, *36*, 929–964.

[7]Farrington, D.P. (1998). Predictors, causes and correlates of male youth violence. In M. Tonry & M.H. Moore (Eds.), *Youth violence: Vol. 24*. (pp. 421-447). Chicago: University of Chicago Press.

[8]Farrington, D.P., Barnes, G.C., & Lambert, S. (1996). The concentration of offending in families. *Legal and Criminological Psychology, 1*, 47–63.

[9]Farrington, D.P., & Hawkins, J.D. (1991). Predicting participation, early onset, and later persistence in officially recorded offending. *Criminal Behavior and Mental Health*, *1*, 1–33.

[10]Farrington, D.P., Loeber, R., & Van Kammen, W.B. (1990). Long-term universal outcomes of hyperactivity-impulsivity-attention deficit and conduct problems in childhood. In L.N. Robins & M. Rutter (Eds.), *Straight and devious pathways from childhood to adulthood* (pp. 62-81). Cambridge, England: Cambridge Univ. Press.

[11]Sampson, R., & Laub, J. (1993). *Crime in the making: Pathways and turning points through life*. Cambridge, MA: Harvard University Press.

[12]Siegel, L.J. (1998). *Criminology: Theories, patterns, and typologies* (6th ed.). Belmont, CA: Wadsworth Publishing Company.

[13]Curry, G.D., & Spergel, I. (1988). Gang homicide, delinquency, and community. *Criminology*, *26*, 381-407.

[14]Park, R. (1915). The city: Suggestions for the investigation of behavior in the city environment. *American Journal of Sociology*, *20*, 579-583.

[15]Park, R., Burgess, E., & McKenzie, R. (1925). *The city*. Chicago: University of Chicago Press.

[16]Shaw, C.R., & McKay, H.D. (1972). *Juvenile delinquency and urban areas* (Rev. ed.). Chicago: University of Chicago Press.

[17]Cohen, A. (1955). *Delinquent boys*. New York: Free Press.

[18]Cloward, R., & Ohlin, L. (1960). *Delinquency and opportunity*. New York: Free Press.

[19]Bandura, A. (1973). *Aggression: A social learning analysis*. Englewood Cliffs, NJ: Prentice Hall.

[20]Bandura, A. (1977). *Social learning theory*. Englewood Cliffs, NJ: Prentice-Hall.

[21]Bandura, A. (1979). The social learning perspective: Mechanisms of aggression. In H. Toch

2155-0182/10/$3.00

(Ed.). *Psychology of crime and criminal justice* (pp. 198-236). NY: Holt, Rinehart, and Winston.

[22]Sutherland, E. (1939). *Principles of criminology*. Philadelphia: Lippincott.

[23]Sutherland, E., & Cressey, D. (1970). *Criminology* (8th ed.). Philadelphia: Lippincott.

[24]Elliott, D., Huizinga, D., & Ageton, S. (1985). *Explaining delinquency and drug use*. Newbury Park, CA: Sage Publications.

[25]O'Donnell, J., Hawkins, J. D., & Abbott, R. (1995). Predicting serious delinquency and substance use among aggressive boys. *Journal of Consulting and Clinical Psychology*, *63*, 529-537.

[26]Thornberry, T. (1987). Toward an interactional theory of delinquency. *Criminology*, *25*, 863-891.

[27]Thornberry, T., Lizotte, A., Krohn, M., Farnworth, M., & Jang, S.J. (1994). Delinquent peers, beliefs, and delinquent behavior: A longitudinal test of interactional theory. *Criminology*, *32*, 601-637.

[28]Kohlberg, L. (1969). *Stages in the development of moral thought and action*. New York: Holt, Rinehart, and Winston.

[29]Henggeler, S. (1989). *Delinquency in adolescence*. Newbury Park, CA: Sage.

[30]Kohlberg, L., Kauffman, K., Scharf, P., & Hickey, J. (1973). *The just community approach in corrections: A manual*. Niantic, CT: Connecticut Department of Corrections.

[31]Menard, S., & Elliott, D. (1994). Delinquent bonding, moral beliefs, and illegal behavior: A three wave-panel model. *Justice Quarterly*, *11*, 173-188.

[32]Hirschi, T. (1969). *Causes of delinquency*. Berkeley, CA: University of California Press.

[33]Gottfredson, M., & Hirschi, T. (1990). *A general theory of crime*. Stanford, CA: Stanford University Press.

[34]Erickson, K. (1962). Notes on the sociology of deviance. *Social Problems*, *9*, 397-414.

[35]Schur, E. (1972). *Labeling deviant behavior*. New York: Harper & Row.

[36]Yarnold, P.R., & Soltysik, R.C. (2005). *Optimal data analysis: A guidebook with software for Windows*. Washington, DC: APA Book Co.

[37]Howell, D.C. (2002). *Statistical methods for psychology* (5th ed.). Pacific Grove, CA: Duxbury, Thomson Learning.

[38]Yarnold, P.R., Soltysik, R.C., & Bennett, C.L. (1997). Predicting in-hospital mortality of patients with AIDS-related pneumocystis carinii pneumonia: An example of hierarchically optimal classification tree analysis. *Statistics in Medicine*, *16*, 1451-1463.

[39]Bryant, F.B. (2005). How to make the best of your data. [Review of the book *Optimal data analysis: A guidebook with software for Windows*] [Electronic version] *Contemporary Psychology: APA Review of Books*, *50*.

[40]Elliott, D. (1977). *National youth survey [United States]: Wave I, 1976* [Computer file] ICPSR version. Boulder, CO: University of CO, Behavioral Research Institute [producer]. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 1994.

[41]Elliott, D. (1986). *National youth survey [United States]: Wave III, 1978* [Computer file] ICPSR version. Boulder, CO: University of CO, Behavioral Research Institute [producer]. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 1994.

[42]Johnston, L.D., O'Malley, P.M., Bachman, J.G., & Schulenberg, J.E. (December 21, 2004). *Overall teen drug use continues gradual de-*

2155-0182/10/$3.00

cline; but use of inhalants rises. University of Michigan News and Information Services: Ann Arbor, MI. [On-line]. Available: http://www.monitoringthefuture.org; accessed 03/16/05.

[43]Henry J. Kaiser Family Foundation. (May 19, 2003). *National survey of adolescents and young adults: Sexual health knowledge, attitudes and experiences.* Menlo Park, CA. [On-line]. Available: http://www.kff.org; accessed 03/16/05.

[44] Ostrander, R.,Weinfurt, K.P., Yarnold, P.R., & August, G. (1998). Diagnosing attention deficit disorders using the BASC and the CBCL: Test and construct validity analyses using optimal discriminant classification trees. *Journal of Consulting and Clinical Psychology*, *66*, 660-672.

[45]Clarke, R. (1995). Situational crime prevention. In M. Tonry (Series Ed.) & D. Farrington (Vol. Ed.), *Building a safer society: Vol. 19. Strategic approaches to crime preventioni (pp. 91-151)*. Chicago: University of Chicago Press.

[46]Lizotte, A., Mercy, J., Monkkonen, E. (1982). Crime and police strength in an urban setting: Chicago, 1947-1970. In J. Hagan (Eds.), *Quantitative criminology*. (pp. 129-148). Beverly Hills, CA: Sage.

Author Notes

Work on this project was completed as a master's thesis by Hideo Suzuki while he was a graduate student in the Department of Psychology, Loyola University Chicago. He is currently a Postdoctoral Research Associate in the Department of Psychiatry, Washington University. Fred B. Bryant and John D. Edwards are faculty in the Department of Psychology, Loyola University Chicago. We are grateful to Paul R. Yarnold and Robert C. Soltysik for advice and guidance in data analysis, and to Delbert S. Elliott, who conducted the National Youth Survey. eMail: suzukih@psychiatry.wustl.edu .