

What works in Offender Profiling? A Comparison of Typological, Thematic, and Multivariate Models

Alasdair M. Goodwill, Ph.D.^{*},
Laurence J. Alison, Ph.D.[†]
and Anthony R. Beech, Ph.D.[‡]

Utilizing a sample of 85 stranger rapists, three models (Hazelwood's (1987) *Power and Anger* FBI model, the *Behavioral Thematic* evaluation of Canter, Bennell, Alison, and Reddy (2003), and the Massachusetts Treatment Center: Rape classification system revision 3 (MTC:R3, Knight & Prentky, 1990)) were contrasted with a *multivariate* regression approach to assess their ability to predict an offender's previous convictions from crime scene information. In respect of the three aforementioned models, logistic regression and AUC analysis indicated that the Power and Anger FBI model was the most effective, followed by the MTC:R3, and then the Behavioral Thematic evaluation. However, predictive analyses based on a multivariate approach using a mixture of crime scene behaviors, as opposed to the grouping of behaviors into themes or types as in the three models, far exceeded the predictive ability of the three models under AUC analysis. The results suggest that emphasis should be placed on further exploration of the predictive validity of each of the individual behaviors that comprise existing thematic, typological, and multivariate classification systems, especially those that are subject to inter-situational variation. Copyright © 2009 John Wiley & Sons, Ltd.

INTRODUCTION

In the period between 2002 and 2005, there was a 41% increase in the number of rapes reported to the police in the U.K. During the same period, the detection rate

*Correspondence to: Alasdair M. Goodwill, Ph.D., Centre for Forensic and Criminological Psychology, University of Birmingham, School of Psychology, Edgbaston, B15 2TT, UK.

E-mail: a.m.goodwill@bham.ac.uk

[†]University of Liverpool, U.K.

[‡]University of Birmingham, U.K.

declined from 41 to 30% (HMcpSI, 2007). The introduction of legislation broadening the definition of rape to include male victims/anal and oral penetration (Criminal Justice and Public Order Act, 1994, and Sexual Offences Act, 2003, respectively) may partially account for increased report rates, but within the political aims of tackling fear of crime, as well as crime itself, such figures constitute an important policy issue, especially when the media report rape *convictions* of below 6% and contrast these figures with 90% *detection* rates for murder (BBC News, February 25, 2005).

Although not confined, by any means, to sexual offenses, the process of 'offender profiling'¹ is perhaps most popularly conceived of as associated with serious sexual offenses and has attracted considerable media and academic attention in recent years (see Ainsworth, 2001; Alison & Canter, 1999; Canter, 2000, for reviews), because such behavioral advice is often used to assist or guide police investigations. However, careful examination of profile advice is warranted, as profilers' advice can have serious consequences in the investigation of serious offenses, because ineffective profiling at the investigative stage can have enormous costs in terms of delivery of justice as well as potentially wasting limited public resources. Hence, in the long term, its success or otherwise may impact on policing practice and policy.

Profiling as an activity has emerged from scientific, technical, and other specialized knowledge, and in some countries profiling may be admissible as evidence in court, for example, under *Daubert* criteria in the U.S. (113 S. Ct. 2786 [1993]). The four *Daubert* criteria for evaluating the admissibility of expert testimony are (1) whether the methods upon which the testimony is based are centered upon a testable hypothesis, (2) the known or potential rate of error associated with the method, (3) whether the method has been subject to peer review, and (4) whether the method is generally accepted in the relevant scientific community. Generally, offender profiling as a process is broadly defined as one in which the profiler infers offender characteristics (e.g. prior convictions, age, employment) from crime scene information (e.g. victim age, sexual acts, level of aggression) (Ault & Reese, 1980; Douglas, Burgess, Burgess, & Ressler, 1992; Hazelwood & Burgess, 1987; Wilson, Lincoln, & Kocsis, 1997). Canter (2000) specifically identifies this as the so-called A-C equation (actions-characteristics), in which he indicates that the object is to identify which crime scene behaviors (actions), or clusters thereof, predict which offender characteristics or clusters of characteristics.

Profiling Approaches and Models

The range of profiling techniques is widespread, with each expert coming from different perspectives and backgrounds, often utilizing different psychological and investigative models in their provision of profiling and investigative advice (Homant & Kennedy, 1998; Muller, 2000; Wilson *et al.*, 1997). Three such models² commonly used in the provision of profiling advice of sexual offenders will be the focus of the current study: they are Hazelwood's (1987) *Power and Anger* model used

¹ Also known as psychological profiling, criminal profiling, and behavioral investigative advice.

² The three approaches will be referred to as 'models' and collectively as 'the models' throughout this manuscript. The multivariate approach defined later will not be referred to as a 'model' nor included in the collective term 'the models'.

by the American Federal Bureau of Investigation (FBI); the *Behavioral Thematic Evaluation* of Canter et al. (2003); and the Massachusetts Treatment Center: Rape Classification system (MTC:R3, Knight & Prentky, 1990).

Hazelwood's Power and Anger model, broadly representative of a criminal investigative approach (National Center for the Analysis of Violent Crime, 1990) is employed by law enforcement agencies, such as the FBI, to aid in classifying rapists based on investigative experience and expertise. Early criminal investigative analysis involved the classification of sexual homicides (e.g. lust murderers) according to how *organized* or *disorganized* the behavior at a crime scene was, with the implication that this would relate to the offender's everyday life and behavior as reflected in his characteristics (Hazelwood & Douglas, 1980). Later, Hazelwood (1987) refined this dichotomy by drawing heavily on the work of Groth, Burgess, and Holmstrom (1977) and Groth and Birnbaum (1979), to create a five component typology for profiling rapists. The resulting typology consisted of the four types of rapist of Groth et al.: (1) *power-reassurance*—offender doubts sexual adequacy and masculinity; (2) *power-assertive*—offense expresses virility, mastery, and dominance; (3) *anger-retaliatory*—offense expresses hostility and rage against women; (4) *anger-excitation*—offense expresses eroticized sadistic aggression, with Hazelwood's addition of a (5) *opportunistic* rapist who commits rape in conjunction with another crime in which rape was not the principal reason for the offense (e.g. burglary).

One of the earliest academic-based typologies for profiling sex offenders was the Canter and Heritage (1990) investigative psychological approach that suggested that offenders could be classified, distinct from inferred motives, into five typological categories; *intimacy*, *sexuality*, *violence*, *impersonal*, and *criminality*. Later, Canter (1994) analyzed 105 cases of rape and concluded that rapist behavior could be defined in terms of behavioral theme or role (i.e. person, victim, or vehicle) that the victim plays for the offender. More recently, Canter et al. (2003) have suggested that stranger rapes can be classified not only by behavioral themes and roles but also by the severity and type of victim violation (i.e., personal, physical, and sexual). On this basis, Canter et al. suggested that rapists could be classified into one of four thematic styles: (1) *hostility*—where the offender uses aggression and violence to demean or humiliate the victim; (2) *control*—where the offender uses controlling behaviors to immobilize the victim; (3) *theft*—where the offender uses the opportunity for some instrumental goal (e.g., financial gain); (4) *involvement*—the offender attempts some form of pseudo-intimacy with the victim. However, Canter et al. found that 27% of the offenses could not be classified by this academic typology and therefore an additional hybrid, or (5) *mixed* group, was added.

The Massachusetts Treatment Center Rape Classification System (MTC:R1) developed by Prentky, Cohen, and Seghorn (1985), currently in its third revision (MTC:R3; Knight and Prentky, 1990), is one of the most rigorously tested classification systems in sexual offending research (Knight, 1999; Knight & Prentky, 1990). The principal aim of the MTC:R3 is as a classification system for research, assessment, and treatment of sex offenders. However, research by Knight, Warren, Reboussin, and Soley (1998) attempted to predict rapist types from crime scene variables to potentially aid in police investigations, reporting some success in the prediction of antisocial and expressive aggression types.

The MTC:R3 consists of five major rapist types: (1) *opportunistic*—where the rape is an impulsive act and is situationally driven (e.g., the rape occurs during the

commission of another crime); (2) *pervasive anger*—where the rape is part of global and undifferentiated anger towards others; (3) *sexual sadistic*—where offense is driven by a fusion of sexual and aggressive feelings; (4) *sexual (non-sadistic)*—where the offender has an enduring sexual preoccupation and feelings of inadequacy; and (5) *vindictive*—where an offender's motivation involves exclusive misogynistic anger.

Selection of Profiling Models

These models were selected as representative of a cross-section of offender profiling research because they have (1) similar thematic underpinnings, (2) analytical feasibility, and (3) previous utility in profiling offender preconvictions.

The three models have similar thematic underpinnings, namely aggression, sexualization, and impulsivity. This is perhaps unsurprising in terms of Hazelwood's Power and Anger model and the MTC:R3, as both were derived from the early work of Murray Cohen, Theoharis Seghorn, and Nicholas Groth in the 1960s and 1970s. Hazelwood's Power and Anger model is an extension of the model of Groth *et al.* (1977), simply adding an opportunistic offender to the four rapist types of Groth *et al.*. However, the MTC:R3 is a more empirically based and validated revision of the original Cohen and Seghorn system, the MTC:R1, which itself was developed in parallel with the system of Groth *et al.* (1977) (Knight & Prentky, 1990).

The Behavioural Thematic model of Canter *et al.* (2003), although not developed on the basis of the offender motivation models of Groth *et al.* (1977), adopts typological groupings that correspond agreeably with both the Power and Anger and MTC:R3 models. Therefore, although each model was born utilizing different processes and objectives, the thematic and typological groupings are similar.

The models utilized in this study were selected primarily for their analytical feasibility and transparency for this comparative study. It would no doubt have been of interest to include a clinically based psychological or psychiatric profile (Boon, 1997; Copson, Babcock, Boon, & Britton, 1997; McGrath, 2000) or diagnostic evaluation (see Wilson *et al.*, 1997); however, these approaches are difficult to compare empirically for two reasons. First, the profiles are often based on the motivation, desires, and fantasies of an offender, which are impossible to verify reliably post-conviction (Alison, Smith, Eastman, & Rainbow, 2003). Second, the approach is often difficult to quantify because of the apparent trade-off between the large number of cases required for quantitative research and the lengthy detailed clinical analysis needed for each individual case (Wilson *et al.*, 1997). This is not to say that clinical-based approaches may or may not be equally valid; simply, in this comparative and quantitative context they were not feasible. Accordingly, only models with published descriptions of typological or thematic categorization (Canter *et al.*, 2003; Hazelwood, 1987; Knight *et al.*, 1998) were selected.

Finally, the Power and Anger (Hazelwood & Warren, 1989), Behavioural Thematic (Hakkanen, Lindlof, & Santtila, 2004) and the MTC:R3 (Knight *et al.*, 1998) models have all been used previously to predict offender characteristics of stranger rapists. Therefore, some precedent in the academic literature for using these models for profiling offender characteristics has already been established. However,

there has been no empirical research to investigate how these models compare with one another in predicting offender preconvictions.

The distinction between these three models represents a difference in the theoretical perspectives and background training of individuals in these areas. Yet, all have been applied to the practical application of offender profiling with little understanding of the comparative advantages and disadvantages of the disparate models. Moreover, researchers have differed significantly in what crime scene information (e.g. behaviors) they believe should be used, what offender characteristics can and should be predicted, and how this causal link between crime scene actions and offender characteristics should be made (see Wilson et al., 1997, for a review). The present study acknowledges the lack of any effort in the field to provide a systematic or comprehensive analytical backdrop to the prediction of offender characteristics from crime scene information. Therefore, one of the aims of the present study was to provide analysis for a range of models and approaches in the prediction of a broad array of possible offender preconvictions. Clearly, the identification *and* predictive ability of even the most basic underlying factors in the determination of offender characteristics has in the most part lacked empirical investigation.

The Multivariate Profiling Approach

As far as we are aware, there are no systematic studies that have sought to compare these models in the context of predicting offender characteristics, nor have studies considered the potential added advantage of the process of combining clusters of behaviors (e.g. themes and types) over the more simple bivariate relationship between individual behaviors and background characteristics. Behavior may vary widely across situations; for example, the same offender may gag only those victims who scream. Hence, a behavior specific to some situations but not others is not necessarily reason to reject a whole typological or thematic approach. The offender characteristics may still fit broadly within a theme or particular typology and the offenses may still be linked.

It is nonetheless useful, as Alison, West, and Goodwill (2004) previously suggested, to carefully examine the individual behaviors selected for analysis and hence included in the themes or types. Alison, Bennell, Mokros, and Ormerod (2002) have also emphasized this point by analogy to the consideration of traits and situational variation, noting that selection of specific variables for analysis is critical in discriminating those elements that are likely to be a product of individual differences compared with those that may be context specific. Canter (2000) has also indicated the importance of establishing the individual 'weighting' of specific variables, and he and his colleagues have conducted several studies in which one can compare the replicability of the themes found in rape (Alison & Stein, 2001, Canter et al., 2003; Canter & Heritage, 1990).

These studies do indicate that there is some behavioral variation across studies and some differences in the interpretations of the underlying themes, suggesting that further work needs to establish which specific behaviors occur with sufficient regularity in meaningful ways. Canter et al. (2003) argue that 'discrete crime scene behaviors may be less significant than the underlying themes of the offence: themes

that can be expressed by the offender in a number of behaviorally different ways' (p. 171).

The present paper will compare the predictive ability of the three models outlined above in order to determine whether any afford added predictive ability in contrast to a multivariate approach. The multivariate approach will investigate the predictive ability of multiple crime scene variables, taken from the Canter *et al.* (2003) model, at once in the prediction of offender preconvictions under full logistical regression, the aim being to advance the understanding of the underlying factors in the classification of offenders, and how these factors may predict offender characteristics in order to prioritize suspects in police investigations. Furthermore, such a pragmatic approach to offender profiling is designed to identify the practical and elementary foundations of offender characteristic prediction.

METHOD

Sample

The sample consists of 85 stranger rapes of single victims committed by 85 lone offenders. For the purpose of this study, a 'stranger rape' is defined as a rape where the offender and victim were not known to one another, or known only to the offender for a brief period of time (e.g. met at a bar a few hours before) prior to the assault. The offenses all took place in large cities within the UK between 1997 and 2002. All 85 offenders have been convicted for the index offense of rape. The sample consists of 57 case records from the English and Welsh Prison Service (i.e., offenders were incarcerated at the time of the study) and 28 cases from Hampshire constabulary, a large police force in the southwest of the U.K. The mean age of the offenders in the sample was 28.6 ($SD = 8.639$). The race of 62 (72.9%) of the offenders was classed as white, 13 (15.3%) African–Caribbean, 6 (7.1%) Asian and 4 (4.7%) were of another or mixed race.

The data used in the study contained detailed descriptions of the cases transcribed from official police, prosecution, and court records. In the majority of cases this also included interviews with the offender in regards to the index offense (Beech, Oliver, Fisher, & Beckett, 2006). Evidence utilized in the present study extracted from these interview transcripts were corroborated against the relevant prison file information. Police data were extracted from computerized police databases and paper records on site. The police data sample contained information about the crime, victim, and offender. This included detailed descriptions of the offense from victim statements.

The 85 stranger rape offenses were analyzed by two independent raters. Training in the coding process involved a three-step process. First, the two raters reviewed the methodology and theory behind each different approach by reading relevant literature on the Power and Anger typology (Hazelwood, 1987; Groth *et al.*, 1977), Knight and Prentky (1990), Knight *et al.* (1998), and Knight (1999) for the MTC:R3 typology, and Canter *et al.* (2003) for specific variable coding. Second, the raters read five mock cases and scored each variable as present or absent under Canter *et al.* methodology and made typological ratings for the MTC:R3 and Power and Anger typologies based on typological descriptors of offender types (available from the authors on request). Feedback was then given and agreement of typological

ratings between raters was reached. Next, the raters reviewed a second set of five mock cases and scored them. The primary researcher again gave feedback on the decisions. The files were then reviewed by the raters. Differences arising in classification were resolved by consensus between raters.

Variable Coding

Canter et al. (2003) Behavioral Thematic and Multivariate Coding

The variables used to code this model were the 27 variables reported by Canter et al. (2003). The inter-rater reliability of the coding of the 85 cases was assessed using Cohen's kappa and interpreted using Fleiss (1981) guidelines. Fleiss suggests that kappas between 0.4 and 0.6 are 'fair', between 0.6 and 0.75 are 'good', and over 0.75 are 'excellent'. The kappa inter-rater reliability between the coders ranged from the 'excellent' (kappa = 0.92, $T = 17.23$, $p < 0.001$) for *vaginal penetration* to 'fair' for *surprise attack* (kappa = 0.53, $T = 5.38$, $p < 0.001$). The average reliability (sum of all variable kappas divided by the number of variables) was in the 'good' range (kappa = 0.71). The highest frequency behavior was *vaginal penetration*, 89.4% (76 offenses), and the lowest frequency behavior was the *offender disguising his identity*, 2.4% (two offenses).

Each offense was then classified into one of the four typological themes of *control*, *hostile*, *involvement*, and *theft* based on the procedure reported by Canter et al. (2003). Here, for example, Canter et al. suggest that if the number of variables within a particular type, for example *hostile*, was greater than the sum of the number of variables in the other three themes (*involvement*, *control*, and *theft*) then the case was coded as primarily *hostile*. In situations where no dominant theme was present, the case was classified into a fifth *mixed* theme as per the procedure of Canter et al. Of the 85 rapes, 40 were classified as predominantly *hostility* (47%), 32 as *mixed* (38%), six as *involvement* (7%), four as *theft* (5%), and three as *control* (3%). As the process of classifying offenders into themes under the Canter et al. (2003) methodology is purely formulaic, reliability of coding in this approach was perfect. However, the overall reliability of this methodology will depend on the reliability of the original variable coding (as above).

MTC:R3 and Power and Anger Model Coding

In respect of the MTC: R3, 13 (15%) offenders were classified as opportunistic, 15 (18%) pervasive anger, 11 sexual sadistic (13%), 27 (32%) sexual non-sadistic, and 19 vindictive (22%). In ratings of cases in regards to the Power and Anger model, 19 (22%) were assigned to the power-reassurance category, 18 (22%) power-assertive, 23 (27%) anger-retaliatory, 10 (12%) anger-excitation, and 15 (18%) opportunistic. The kappa inter-rater reliability scores between the two coders for the MTC:R3 and Power and Anger typologies (kappa = 0.84, $T = 15.04$, $p < 0.001$ and kappa = 0.82, $T = 14.69$, $p < 0.001$, respectively) were 'excellent' (Fleiss, 1981).

Offender Preconvictions

Offender preconvictions were extracted from police and prison case files and cross-referenced on the U.K. Police National Computer (PNC). Offenders were considered to have a preconviction for a certain type of offense if they had either a juvenile or adult criminal record for that offense type. For analytical simplicity offender preconvictions were categorized into five mutually exclusive categories; *sexual*, *property*, *violent*, *drug and/or weapon*, and *other* offenses. Of the sample, 38 (44.7%) offenders had preconvictions for *sexual* (e.g. gross indecency, buggery, rape, indecent exposure, indecent assault), 57 (67.1%) had *property* (e.g. fraud, criminal damage, theft, burglary, vehicle theft), 48 (56.5%) had *violent* (e.g. assault, harassment, affray, battery, robbery, public disorder, actual bodily harm, grievous bodily harm), 22 (25.9%) had *drug and/or weapon offenses*, and 44 (51.8%) had *other* offenses (e.g. driving disqualified, public disorder, failure to appear). Overall, 71 (83.5%) of the offenders had at least one adult or juvenile preconviction.

Analysis

Logistic regression analysis was used to examine the ability of each of the three models, Power and Anger, MTC:R3, and Behavioural Thematic, to predict offender preconvictions. In addition, the multivariate model (e.g. the 27 variables of the Canter *et al.* typology) was regressed in a full model (enter method: SPSS for Windows, Ver. 14.0) in order to predict each of the preconviction categories. The predicted probabilities of categorical membership (e.g. has the preconviction/does not have the preconviction) for each preconviction prediction for each approach were then entered into separate ROC analyses.

Receiver operating characteristic analysis (Swets, 1988) was used to assess the predictive ability of the models in comparison to one another. ROC analysis utilizes area under the curve (AUC) statistic to assess both the hit probability (pH) and the false alarm probability (pFA) simultaneously, providing a measure for predictive accuracy independent of decision thresholds, unlike typical logistic regression classification tables. The hit probability (pH) or *sensitivity* assesses the probability that the regression model (i.e. the typology) has correctly predicted the presence of the preconviction in question, whereas the false alarm probability (pFA) or $1 - \text{specificity}$ assesses the probability that the model has incorrectly predicted the presence of the preconviction in question.

The AUC analysis graphically plots the relationship between p(H) and p(FA) as a relational curve. The area under the curve (AUC) is therefore a measure of the diagnostic accuracy of the particular prediction, in our case the prediction of a particular preconviction. The AUC can range from 0 (indicating no accuracy) to 1.0 (indicating perfect accuracy), where the midpoint of 0.5 indicates only chance accuracy.

Utilizing ROC analysis and AUC statistics in the prediction of offender characteristics in the present study is advantageous for two reasons. First, the AUC is a measure of the predictive ability of each approach independent of any decision threshold. This is because the AUC is a measure of diagnostic accuracy across all decision thresholds (i.e. the entire curve), not just one particular threshold, as in

percentage correct estimates (Swets, 1988). When utilizing a single decision threshold it is impossible to determine whether the level of accuracy achieved is a result of the specific predictors or simply a result of the decision threshold adopted (Bennell & Jones, 2005).

Second, the AUC statistic is based on the predictive outcomes of each approach rather than their raw frequencies, allowing for comparisons to be made between approaches regardless of the relative frequency of each predictive alternative. This is particularly advantageous in the present study as we are comparing the relative predictive ability of several approaches with varying numbers of predictors.

In the current study, the ROC (AUC) analysis was calculated using MedCalc for Windows, Version 9.2.1.0 (MedCalc Software, Mariakerke, Belgium). The statistical method of comparing ROC curves differs based on whether the ROC curves have been produced from the same or different populations. In our case, statistical comparisons are made on ROC curves from the same population. In these circumstances it is inappropriate to calculate the standard error of the differences between areas, as the areas are likely to be correlated. The MedCalc software utilizes the Hanley and McNeil (1983) method for testing differences between ROC curves produced by correlated (or identical) samples. The approach by Hanley and McNeil takes into account the correlated nature of samples by using the maximum likelihood estimation of Dorfman and Alf (1969) for the AUC and a nonparametric tau to calculate the correlation between the values of the original measures (DeLong, DeLong, & Clarke-Pearson, 1988).

RESULTS

Logistic Regression Models

Table 1 contains a summary of the regression results, including the beta (β), standard error (SE), significance (SIG), model χ^2 and Cox & Snell R^2 values for sexual, property, violent, drugs and/or weapon, and other offenses.

Sexual Preconvictions

Table 1 illustrates that the multivariate model had a non-significant fit to the data and explained 29% of the variance. Only the *anal penetration* variable reached a significant beta value ($\beta = 2.12$, $p < 0.05$) suggesting that when all variables are considered simultaneously anal penetration by the offender is 8.4 times (odds ratio = $\exp(B)$) more likely to occur in offenses in which the offender has a sexual preconviction. The MTC:R3 model was significant ($\chi^2 = 10.22$, $p < .05$) and accounted for 11% of the variance. Offenders who were typed as *pervasively angry* were 6.7 times more likely to have a sexual preconviction than those typed as *opportunistic*, the contrast variable. The Canter et al. and Power and Anger models both were a poor fit to the data, indicated by their non-significant χ^2 model scores and low R^2 (variance) values ($\chi^2 = 6.99$, $R^2 = 7.9\%$ and $\chi^2 = 7.06$, $R^2 = 8.0\%$, respectively). However, offenders typed as *anger excitation* within the Power and Anger model were significantly more likely (6.4 times) to have a sexual preconviction than those who were *opportunistic*, the contrast variable.

Table 1. Logistic regression analysis comparison chart indicating beta value (β), standard error (SE) and significance level of each regression coefficient, overall model chi-square (X^2), significance, and Cox and Snell (R^2) variance estimate

	SEXUAL	PROPERTY	VIOLENT	WEAPON/DRUG	OTHER	OFFENDER AGE
Variable / Category	β SE $Sig.$ Model $Sig.$ R^2 X^2 Y X R^2	β SE $Sig.$ Model $Sig.$ R^2 X^2 Y X R^2	β SE $Sig.$ Model $Sig.$ R^2 X^2 Y X R^2	β SE $Sig.$ Model $Sig.$ R^2 X^2 Y X R^2	β SE $Sig.$ Model $Sig.$ R^2 X^2 Y X R^2	β SE $Sig.$ Model $Sig.$ R^2 X^2 Y X R^2
<i>Multivariate Approach and justification (why different predictors in regression are indicated)</i>	-0.188 1.471 0.898 28.668 0.326 0.286 -1.473 1.737 0.396 29.735 0.279 0.295 2.124* 0.991 0.032	No significant variables	2.872* 1.441 0.046 -3.217* 1.359 0.009 1.516* 0.717 0.054	No significant variables	No significant variables	No significant variables
<i>Behavioral Theme</i>	Constant -0.647 0.372 0.082 7.062 0.113 0.080 -0.047 0.943 0.961 21.85 23205.4 0.999 0.647 1.087 0.545 0.647 0.488 0.186 Indicative 0.725	Constant 0.938* 0.393 0.017 3.974 0.410 0.046 -0.938 0.906 0.301 20.26 23205.4 0.999 0.160 1.220 0.895 -0.427 0.511 0.403 Indicative 0.826	Constant 0.251 0.356 0.481 2.043 0.728 0.024 -0.944 0.956 0.313 0.442 1.276 0.729 0.847 1.208 0.483 0.051 0.479 0.915 Indicative 0.757	Constant -1.099** 0.408 0.007 0.442 0.979 0.005 -0.511 1.169 0.662 0.405 1.291 0.753 0.000 1.225 1.000 0.129 0.540 0.811 Indicative 0.981	Constant -0.379 0.360 0.292 2.976 0.562 0.054 0.379 0.892 0.671 1.073 1.277 0.401 0.379 1.063 0.721 0.785 0.483 0.104 Indicative 0.573	Constant 0.251 0.356 0.481 2.043 0.728 0.024 -0.944 0.956 0.313 0.442 1.276 0.729 0.847 1.208 0.483 0.051 0.479 0.915 Indicative 0.757
<i>MTOM</i>	Constant -1.204 0.658 0.067 10.217* 0.037 0.113 1.309 0.803 0.103 1.897** 0.856 0.027 1.764 0.909 0.052 0.339 0.782 0.665 Indicative 0.052	Constant 0.154 0.556 0.782 6.224 0.183 0.071 -0.470 0.570 0.410 3.715 0.444 0.043 0.619 0.744 0.405 -0.021 0.760 0.978 2.148 1.187 0.070 0.711 0.698 0.208 Indicative 0.332	Constant -0.470 0.570 0.410 3.715 0.444 0.043 0.575 0.732 0.432 1.163 0.791 0.141 1.431 0.885 0.101 0.693 0.689 0.315 Indicative 0.475	Constant -1.204 0.658 0.067 1.407 0.843 0.016 -0.118 0.866 0.892 0.192 0.800 0.827 -0.300 1.022 0.769 0.511 0.775 0.310 Indicative 0.846	Constant 0.154 0.556 0.782 6.224 0.183 0.071 -0.693 0.732 0.344 -0.288 0.760 0.705 1.309 0.939 0.139 -0.080 0.677 0.906 Indicative 0.272	Constant 1.705* 0.769 0.027 6.390 0.172 0.072 -1.166 0.904 0.197 -1.571 0.927 0.090 -1.887 0.979 0.054 -1.779* 0.860 0.039 Indicative 0.266
<i>Power and Aggr</i>	Constant -1.012 0.584 0.083 6.995 0.116 0.079 0.693 0.746 0.355 0.318 0.769 0.679 1.274 0.720 0.077 1.859* 0.904 0.040 Indicative 0.166	Constant 1.386* 0.645 0.032 12.587* 0.014 0.138 -0.615 0.813 0.451 -1.386 0.799 0.083 -1.124 0.770 0.145 19.82 12710.1 0.999 Indicative 0.439	Constant 0.693 0.548 0.206 4.445 0.349 0.051 -0.588 0.715 0.411 -0.822 0.725 0.206 -0.606 0.689 0.379 0.693 0.962 0.471 Indicative 0.412	Constant -0.134 0.518 0.796 4.393 0.353 0.050 -1.188 0.765 0.120 -0.916 0.725 0.206 -1.425 0.755 0.059 -1.253 0.945 0.185 Indicative 0.354	Constant 0.693 0.548 0.206 12.141* 0.016 0.133 -0.588 0.715 0.411 -1.386 0.742 0.062 -1.135 0.695 0.102 1.504 1.188 0.205 Indicative 0.062	Constant 0.403 0.527 0.442 1.641 0.801 0.019 0.568 0.722 0.011 -0.405 0.707 0.566 -0.143 0.674 0.832 -0.405 0.823 0.622 Indicative 0.809

Property Preconvictions

Table 1 illustrates that multivariate model had a non-significant fit to the data and explained 29% of the variance. The Power and Anger model was significant ($\chi^2 = 12.58$, $p < 0.05$) and explained an adequate 13.8% of the variance, suggesting a good model fit to the data. The Behavioral Thematic model and the MTC:R3 models were not significant and explained little of the variance ($R^2 = 4.6\%$ and 6.7% , respectively).

Violent Preconvictions

Table 1 illustrates that the multivariate model had a non-significant fit to the data and explained 30.4% of the variance. Offenders who *tear clothing* (of the victims) and/or use a *weapon* were significantly more likely in relation to the other variables entered into the model to indicate that the offender had a violent preconviction (17.6 and 4.6 times more likely, respectively). Offenders who performed *cunnilingus* on the victim were 0.04 times (or 96%) less likely to have a preconviction for a violent offense. The Behavioral Thematic model, MTC:R3, and Power and Anger models were all non-significant and explained little of the variance ($R^2 = 2.4\%$, 4.3% , and 5.1% , respectively).

Drugs and/or Weapon Preconvictions

Table 1 indicates that the multivariate model had a near significant ($\chi^2 = 38.64$, $p = 0.053$) fit to the data and explained 36.5% of the variance. In respect of the other variables entered into the model simultaneously, offenders who used a *weapon* were significantly (7.5 times) more likely to have a violent preconviction. The Behavioral Thematic model, MTC:R3, and Power and Anger models all had poor fits (χ^2 significant levels $> 35\%$) and explained little of the variance ($R^2 = 0.5\%$, 1.6% , and 5.0% , respectively).

Other Preconvictions

Table 1 illustrates that the multivariate model had a near significant fit to the data ($\chi^2 = 23.84$, $p = 0.59$ and explained 24.5% of the variance. The Power and Anger model was significant ($\chi^2 = 12.14$, $p < 0.05$) and explained 13.3% of the variance. Offenders typed as power assertive were approximately 0.25 times (75%) less likely to have a preconviction for other offenses, although this value was only significant at the 6.2% level ($p = 0.062$).

ROC (AUC) Analysis

The following sections utilize ROC curves to illustrate the ability of each of the four models to predict each preconviction category and the offender's age. A comparison of each approach's area under the curve (AUC) statistics for each preconviction is

tabulated for each section. Overall, the multivariate approach distinguishes itself as a stronger more accurate model in relation to the preconviction and offender age categories under examination. The Power and Anger and MTC:R3 models both appear to have greater AUC scores than the Behavioral Thematic model across preconvictions; however, the AUC mean differences between the four models were generally found to be non-significant.

Sexual Preconvictions

Figure 1 illustrates the relationship between the models under ROC analysis for *sexual* preconviction prediction. The predictive accuracy of the multivariate model (AUC = 0.81, CI = 0.71–0.89, $p < .001$) was highly significant. Additionally, both the MTC:R3 model (AUC = 0.69, CI = 0.58–0.78, $p < 0.05$), and the Power and Anger model (AUC = 0.66, CI = 0.55–0.76, $p < 0.05$) were significantly accurate in predictions of *sexual* preconvictions.

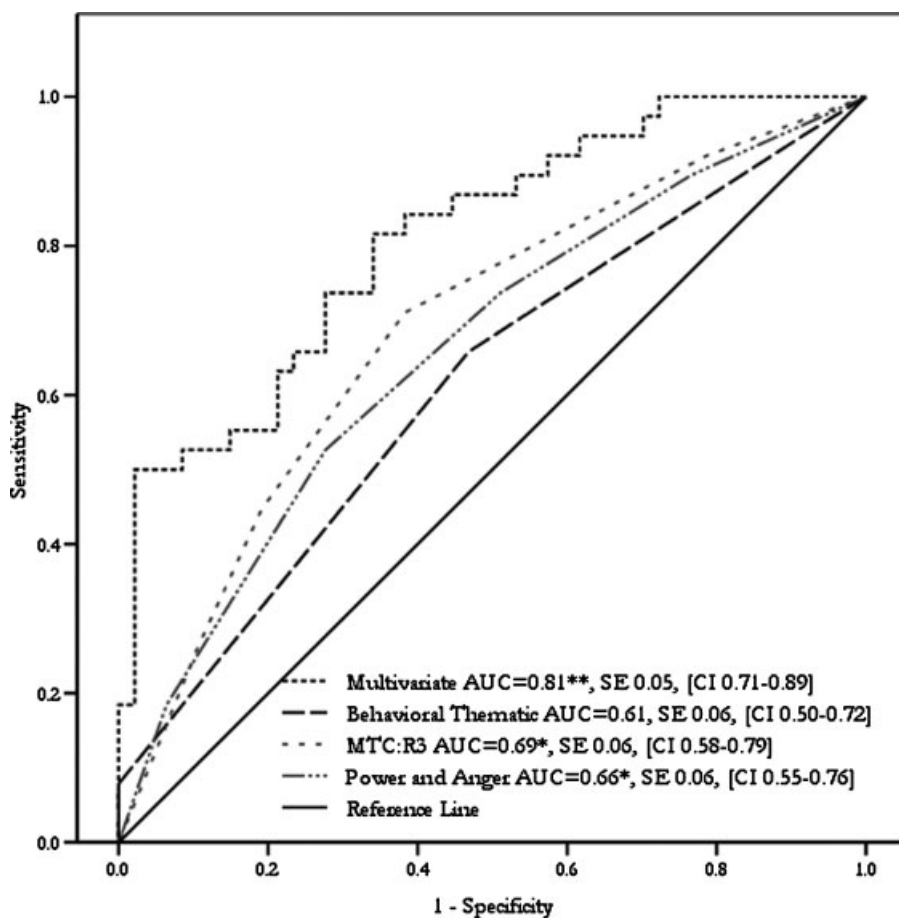


Figure 1. ROC graph of model predictive accuracy for sexual preconvictions (AUC score, standard error (SE) and 95% confidence intervals reported). *Significant at 0.05 level. **Significant at 0.05 level.

A comparison of each model's area under the curve (AUC) statistic revealed that the multivariate model was significantly more accurate than both the Behavioral Thematic model (AUC difference = 0.20, CI = 0.06–0.33, $p < 0.05$) and the Power and Anger model (AUC difference = 0.15, CI = 0.01–0.29, $p < 0.05$) but not the MTC:R3 (AUC difference = 0.12, CI = –0.01–0.26, ns) in prediction of *sexual* preconvictions. The AUC of the MTC:R3 was not significantly different from that of any of the other models.

Property Preconvictions

Figure 2 illustrates the relationship between the models under ROC analysis for *property* preconviction prediction. The predictive accuracy of the multivariate model (AUC = 0.82, CI = 0.72–0.90, $p < 0.001$) was highly significantly different from chance (AUC = 0.5). Additionally, both the MTC:R3 model (AUC = 0.64, CI = 0.53–0.74, $p < 0.05$) and the Power and Anger model (AUC = 0.69, CI = 0.58–0.79, $p < 0.05$)

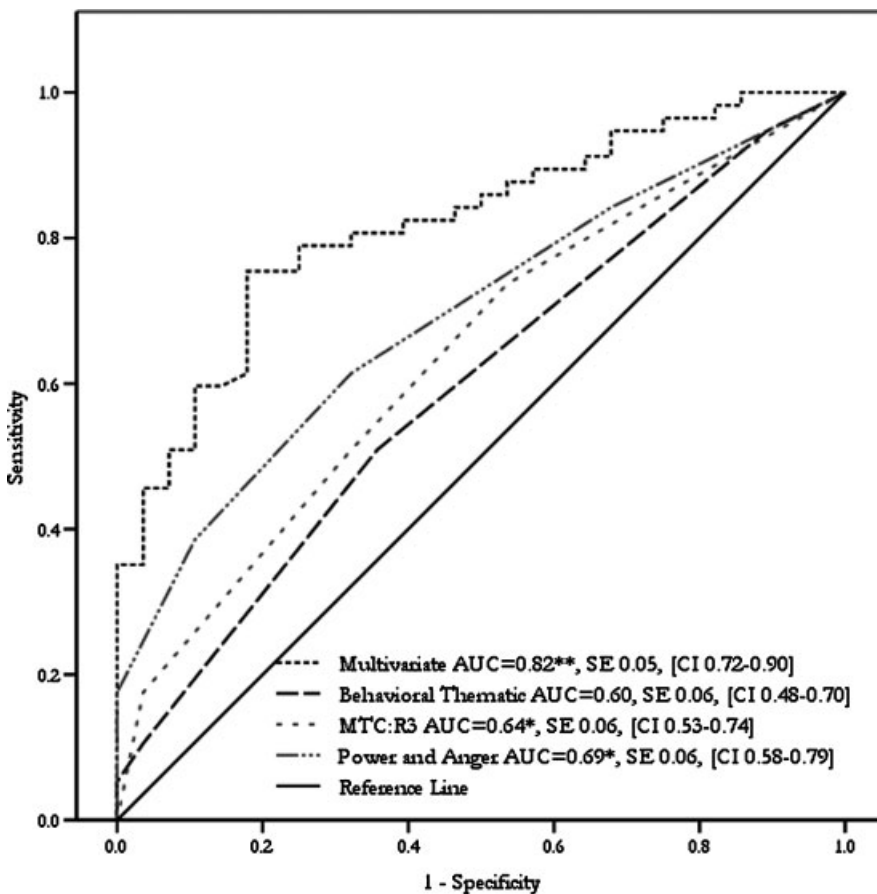


Figure 2. ROC graph of model predictive accuracy for property preconvictions (AUC score, standard error (SE) and 95% confidence intervals reported). *Significant at 0.05 level. **Significant at 0.05 level.

CI = 0.58–0.80, $p < 0.05$) were significantly accurate in predictions of *property* preconvictions.

Comparing the AUC of each model revealed that the multivariate model was significantly more accurate than both the Behavioral Thematic model (AUC difference = 0.22, CI = 0.09–0.36, $p < 0.001$) and the MTC:R3 model (AUC difference = 0.18, CI = 0.04–0.32, $p < 0.05$) but not the Power and Anger model (AUC difference = 0.13, CI = –0.01–0.27, ns) in predicting *property* preconvictions. The Power and Anger model was not significantly different from any of the other models.

Violent Preconvictions

Figure 3 illustrates the relationship between the models under ROC analysis for *violent* preconviction prediction. The predictive accuracy of the multivariate model (AUC = 0.83, CI = 0.73–0.90, $p < 0.001$) was highly significant. The predictive

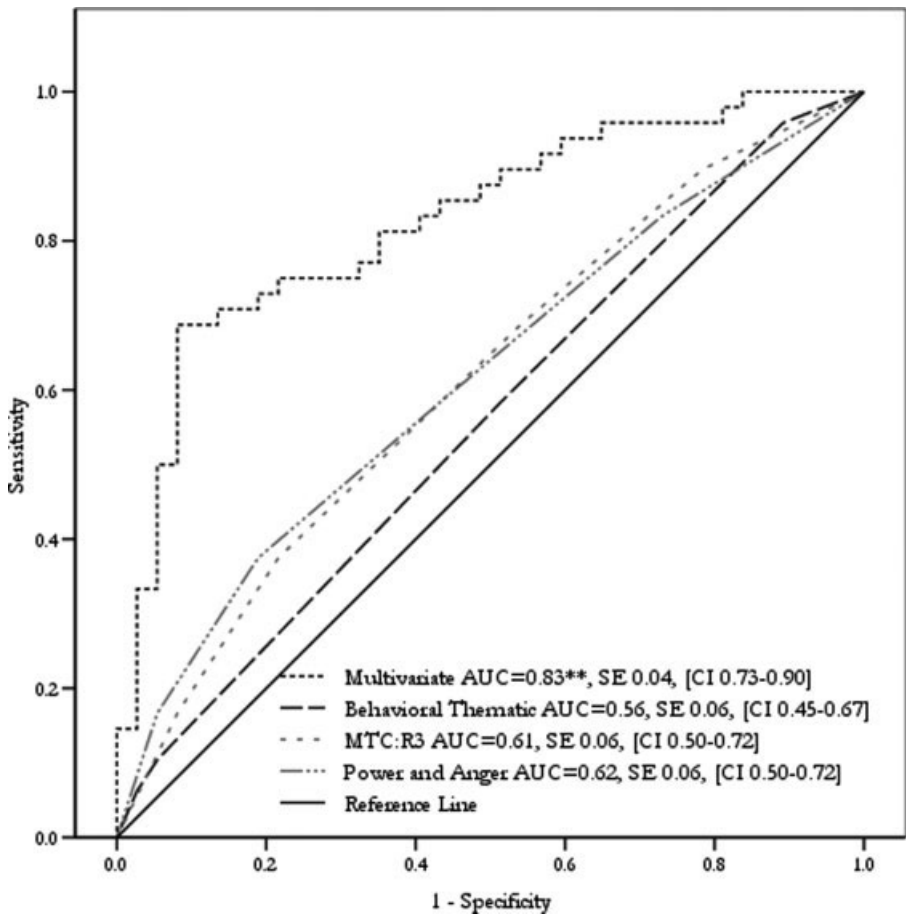


Figure 3. ROC graph of model predictive accuracy for violent preconvictions (AUC score, standard error (SE) and 95% confidence intervals reported). *Significant at 0.05 level. **Significant at 0.05 level.

accuracies for both the MTC:R3 model (AUC = 0.61, CI = 0.50–0.72, $p = 0.08$) and the Power and Anger model (AUC = 0.62, CI = 0.50–0.72, $p = 0.07$) approached significance.

The multivariate model was significantly more accurate than the Behavioral Thematic model (AUC difference = 0.27, CI = 0.13–0.41, $p < 0.001$), MTC:R3 model (AUC difference = 0.22, CI = 0.078–0.353, $p < 0.05$) and the Power and Anger model (AUC difference = 0.21, CI = 0.08–0.34, $p < 0.05$) at predicting *violent* preconvictions. No other models were significantly different from each other.

Weapon and/or Drug Preconvictions

Figure 4 illustrates the relationship between the models under ROC analysis for *weapon and/or drug* preconviction prediction. The predictive accuracy of the multivariate model (AUC = 0.89, CI = 0.80–0.95, $p < 0.001$) was highly significant.

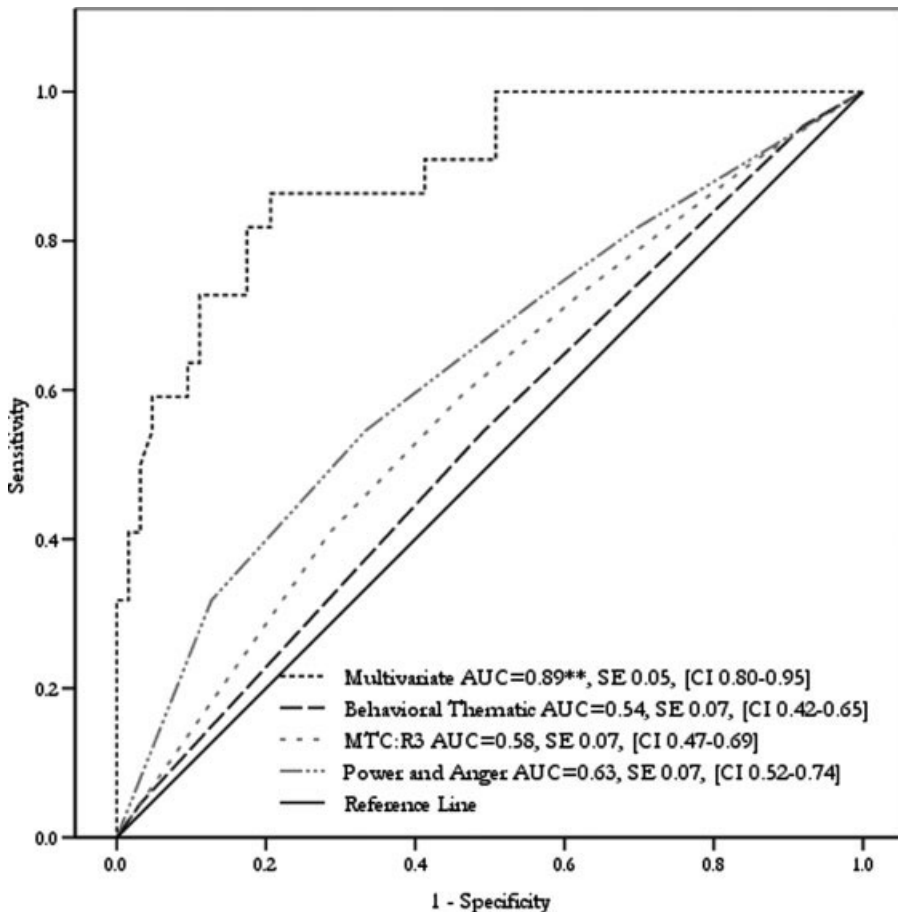


Figure 4. ROC graph of model predictive accuracy for drug and/or weapon preconvictions (AUC score, standard error (SE) and 95% confidence intervals reported). *Significant at 0.05 level. **Significant at 0.05 level.

The predictive accuracy of the Power and Anger model (AUC = 0.63, CI = 0.52–0.74, $p = 0.06$) approached significance.

The multivariate model was found to be significantly more accurate than the Behavioral Thematic model (AUC difference = 0.35, CI = 0.19–0.51, $p < 0.001$), MTC:R3 model (AUC difference = 0.31, CI = 0.15–0.46, $p < 0.001$) and the Power and Anger model (AUC difference = 0.25, CI = 0.10–0.41, $p = 0.001$) under AUC comparison of *weapon and/or drug* offense preconviction prediction. No other models were significantly different from each other.

Other Preconvictions

Figure 5 illustrates the relationship between the models under ROC analysis for *other* preconviction prediction. The predictive accuracy of the multivariate model (AUC = 0.77, CI = 0.63–0.85, $p < 0.001$) was highly significant. Additionally, both the MTC:R3 model (AUC = 0.63, CI = 0.52–0.73, $p < 0.05$) and the Power and

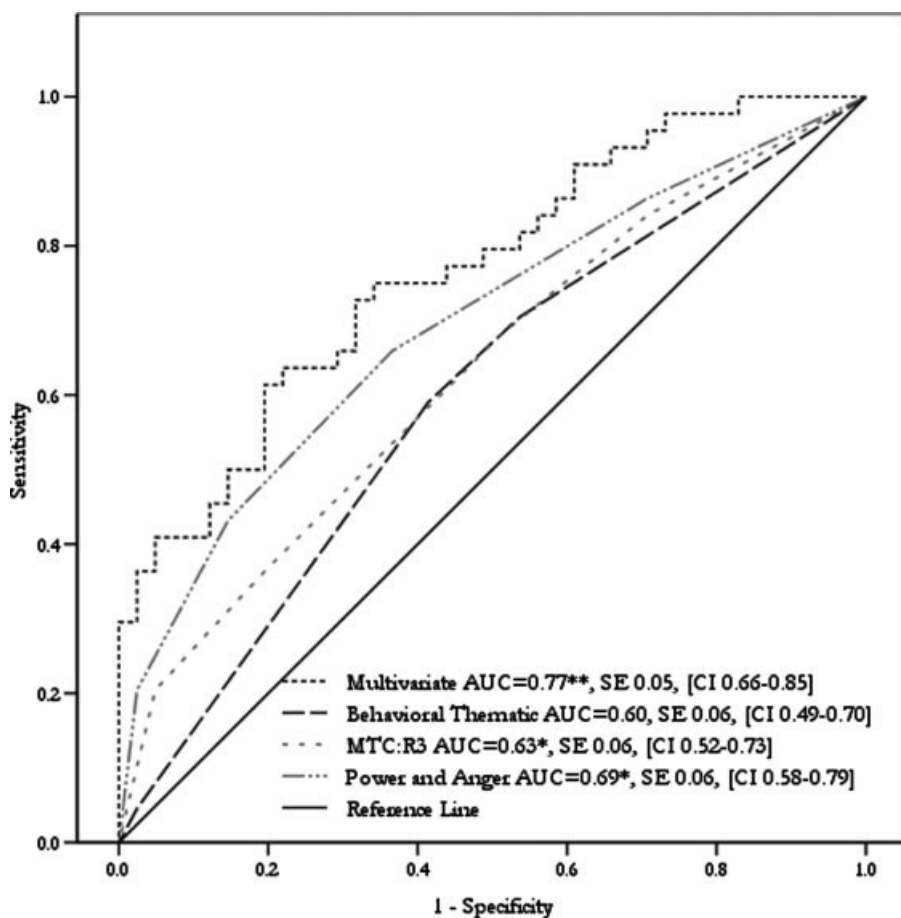


Figure 5. ROC graph of model predictive accuracy for other preconvictions (AUC score, standard error (SE) and 95% confidence intervals reported). *Significant at 0.05 level. **Significant at 0.05 level.

Anger model (AUC = 0.69, CI = 0.58–0.79, $p < 0.05$) were significantly accurate in predictions of other preconvictions.

A comparison of each model's AUC scores revealed that the multivariate model was significantly more accurate than the Behavioral Thematic model (AUC difference = 0.17, CI = 0.02–0.32, $p < 0.05$). No other models were significantly different from each other.

DISCUSSION

Offender profiling is an inexact 'science', but given the implications when such behavioral advice is adopted by police investigators as well as the possibility that such advice may, in some countries, be admissible as evidence in court, it is an important topic for research. In the current research study, three models born from different areas within offender profiling research and practice were contrasted with a simple multivariate approach to examine their ability to predict offender preconvictions. According to the ROC analyses the multivariate logistic approach using the 27 variables of Canter et al. (2003) performed best, followed by the Power and Anger, MTC:R3, and Behavioral Thematic models. Clearly, ranking each approach is somewhat subjective, as multiple tests have been performed with differing outcomes across them. Therefore, this is only provided for discussion purposes as a general indication of the comparative ability of these methods under this specific methodology.

Application to Offender Profiling

Odds ratios give a statistical probability of a certain type of offender having a certain type of preconviction and, thus, have direct utility for criminal investigators. Profilers can ensure that advice is evidence based and, thus, avoid some of the erstwhile criticisms leveled at profiling, namely, that much activity is unscientific and tends to be intuitive or downright idiosyncratic (see Muller, 2000; Hicks & Sales, 2006). In this study, the odds ratios indicated some significant trends across typologies, variables, and preconviction outcome variables. For example, the Power and Anger model regression analyses revealed that the odds of an anger–excitation or anger–retaliatory rapist having a preconviction for a sexual offense were 6.4 times and 3.6 times higher than for an opportunistic offender. Those classified as pervasively angry or sexual sadistic were 6.7 times and 5.8 times more likely to have a prior criminal conviction for a sexual offense than opportunistic offenders under the MTC:R3 classification. Offenders who anally penetrated their victims were also found to be 8.4 times more likely to have a sexual preconviction than those that did not. This identification of offenders displaying a fusion of anger and sexuality in their offenses as a significant predictor of past sexual convictions, in contrast to the myriad of other behaviors that are possible, is an important finding. This is precisely the type of scientific warrant that can aid investigators in prioritizing and refining suspect pools during investigations. For example, when a rape has been overtly anger driven and highly sexual, investigators may prioritize the suspects with prior sexual offenses to be investigated first.

Several significant odds ratios were found within the multivariate approach in the prediction of violent preconvictions. Offenses in which the offender used a weapon (weapon use) or tore the victim's clothing were 17.7 and 4.5 times more likely, respectively, to have a preconviction for a violent offense. On the other hand, offenders who performed cunnilingus on their victims were 25 times less likely to have a violent preconviction. However, no significant odds ratios were identified in any of the typological or thematic models in relation to the prediction of violent preconvictions. This is somewhat surprising, considering that the models all contain clearly identifiable facets of aggression and that past research has identified a link between present violence and a history of violence (see Klassen & O'Connor, 1994, for review).

Other significant findings indicated that offenders with drug and/or weapon preconvictions were 7.5 times more likely to use a weapon than those that did not. Further, anger-retaliatory offenders were four times (significant at the 6% level) less likely than opportunistic offenders to have a drug and/or weapon preconviction.

Impact of Methodological Differences in the Prediction of Offender Preconvictions

Although the offender types classified across the thematic classifications produced less obvious odds ratios and thus arguably more complex relationships with offender preconvictions, they should not in this case be viewed as less effective in terms of odds ratio predictions. Rather, it illustrates a central point in the argument here—namely, that within these typologies and, indeed, subsumed within their subcategories are narrower pools of variables that do successfully predict various preconvictions. For example, many of the discrete variables used within the multivariate approach would have also been considered by raters classifying offenses to types under the MTC:R3 and Power and Anger approaches. Nevertheless, although these behaviors have been considered, their *direct* input into the prediction of offender preconvictions is somewhat lost within all the other contextual and situational aspects of the offense when classification to one type or another is made.

On the other hand, isolating one variable of a multivariate approach to predict an offender characteristic, as in the odds ratio discussions above, somewhat neglects the influence that other information (not included in the analyses) and situational context might have. We were careful to investigate the predictive ability of variables under the multivariate approach in relation to one another by using full logistic models; however, the extrapolation of single significant predictors is inevitable. Goodwill and Alison (2007) recently illustrated that bivariate relationships between crime scene information and offender characteristics can be dependent on the presence or absence of other behavioral, contextual, or situational aspects. Accordingly, Goodwill and Alison (2007) further suggest that the relationship between crime scene information and offender characteristics is likely to be a multifaceted one and interdependent with what other behavioral, contextual, or situational factors are present. This point, circularly, is in fact the premise for the investigation and adoption of thematic and typological approaches: to account for the often found interdependence of behavior and situational complexity.

Hence, it is important to determine not only the most accurate predictors but also the optimal hierarchical level from which to construct profiles. In other words,

should analysis of simple bivariate or multivariate relationships be favored over clusters of behaviors or should sophisticated typologies be adopted? Each have strengths and weaknesses; bivariate (and even multivariate) approaches can neglect the interaction of other behaviors and situational contexts, whereas offenders are often found to be representative across multiple themes or types under thematic evaluations and typologies.

Comparative Ability to Predict Offender Preconvictions

According to the ROC analyses, the multivariate relationship approach was more effective in the prediction of each offender preconviction condition. Having considered only odds ratios and regression results, a significant aspect of predictive analysis would have been missed: the fact that the predictive ability can increase or decrease depending on one's classification cut-off criterion chosen. In logistic regression analysis the typical 0.5 probability cut-off was used to predict whether a preconviction was probable or not; however, this cut-off probability is not necessarily optimal. For example, if we had increased the cut-off criterion from 0.5 to 0.6, we might have caused more offenses to be correctly predicted but likely would have also increased the number of false positives (FP), or false alarms. Thus, an increase in the cut-off criterion is likely to decrease the specificity but increase the sensitivity, whereas a decrease in the cut-off criterion will likely increase the specificity but decrease the sensitivity.

These findings are not merely a matter of methodological or statistical interest to academics. Rather, it is crucial for practitioners to negotiate the best balance between specificity and sensitivity. Basing investigative advice on a cut-off point so high that it generates an unacceptable number of false positives may result in false arrests made by the police, significantly slowing and hindering an investigation. For profiling, the optimal balance between specificity and sensitivity is likely to be a dynamic ongoing enterprise rather than a static one-time decision; it is likely to be dependent on goals and subjective interpretation of the individuals involved.

One explanation for the increased predictive ability of the multivariate approach is in the limited variance accounted for in the other thematic and typological models. Table 1 illustrates this, as the average variance accounted for across offense predictions approximates 30% for the multivariate approach and just 6, 6, and 3% for the MTC:R3, Power and Anger, and Behavioral Thematic models, respectively. The lack of variance explained in the thematic models is likely a result of their rigidity of only five categories with which to classify an offense and predict the outcome measure. Therefore, there seems to be a trade-off between the ability to utilize an array of behaviors as a theme or type to simplify or conceptualize the offense and the ability to generate a sufficient number of categories in the predictor to offer good accountability of variance.

Differences in interpretation could also have potentially created pragmatic limitations. It could be argued that in the Power and Anger and MTC:R3 models offenses were categorized more by the 'feel' of the crime than by their strict 'facts'. In the Power and Anger and MTC:R3 classifications, coders were able to judge offenses as being one type or another, based on their interpretation of the crime scene information, including any situational and contextual factors. In the Behavioral

Thematic and multivariate approaches only the presence or absence of certain behaviors were utilized, with no consideration for situational and contextual influences. For example, consider a hypothetical case where an offender attempts to vaginally penetrate a victim but does not because she convinces him she is menstruating. In utilizing the Behavioral Thematic and multivariate approaches the rater would classify the offense as vaginal penetration being absent, whereas in the MTC:R3 and Power and Anger models coders could consider the situational factors (i.e. victim's verbal deterrent) and judge the behavior accordingly.

The interpretation of case information to classify offenders under the Power and Anger and MTC:R3 models in 'real world' situations will therefore in part be a function of the interpreter's experience, skill, and intuition. Yet, this also suggests that the Power and Anger and less so the MTC:R3 model can be greater influenced by subjectivity and thus potentially more unreliable than the more rigid methodology of the Behavioral Thematic approach. However, this is not necessarily a drawback of these typological models. Given the malleability of these approaches, users *with* experience, skill, and valuable intuition can potentially be more accurate in their classification of offenders precisely *because* the system is less rigid, allowing them to use all available information.

Thematic and typological models, in general, were likely developed to alleviate problems associated with minor ambiguities at the variable level such as these by offering a broader picture of the offending behavior. Interestingly, we have found that a multivariate approach focusing on the specific behaviors in the crime has yielded better predictive accuracy than its typological and thematic counterparts. Nevertheless, it is important to bear in mind that predictive multivariate models are more likely to over-specify bivariate relationships in a data set of which the regression equations have been developed. Therefore, the generalizability of the typological and perhaps more notably the multivariate approach is likely lower in the current analysis than if a split-half, bootstrapping, or jackknifing sampling method had been employed.

Pragmatic Advancement of Offender Profiling

In the current research the Power and Anger typology overall revealed slightly stronger results than the MTC:R3 and in most cases stronger results than the Behavioral Thematic approach. One possible explanation for why this particular typology achieved superior results is that the Power and Anger typological classification is based heavily on the clinical classification of Groth *et al.* (1977) *and* the extensive investigative experience of FBI detectives (Hazelwood, 1987; Hazelwood & Douglas, 1980; Hazelwood & Warren, 1989, 2000). In fact, the Power and Anger typology is more a mixture between the criminal investigative and empirical areas than distinctly falling within one of these areas. It seems that a mixture between disciplines, investigative experience and empirical research, is advantageous in profiling offenders. The current results are indicative of the potential benefits of co-operation and integration that a pragmatic, or mixed, approach has to offer.

Alison *et al.* (2004) argue that a supportive and pragmatic relationship between disciplines is likely to create a far more meaningful solution to the application of

offender profiling. They propose that the creation of profiles should be an integration between academics and practitioners similar to the kind recommended by Fishman (2000) in the therapeutic arena. This proposition is supported by their belief that both camps have different strengths and weaknesses and that 'the sum of the learning of the two groups is potentially much greater than its individual parts' (Alison et al., 2004, p. 72). Effective collaboration across different communities of experts would provide checks and balances and also lend support in meeting Daubert standards for admissibility in court. Nevertheless, such collaboration across domains of expertise must be accompanied by a detailed focus on the typologies themselves, their inherent strengths and weaknesses and, indeed, whether typologies provide the most reliable and valid way forward for profiling to attain the optimum degree of objectivity.

Limitations of the Current Study

It is appreciated that neither the different types of offender profiling research under study, nor the typologies and thematic approaches themselves, define or are representative of all areas of profiling research and application. For example, the MTC:R3, although previously tested for its use as an investigative tool (Knight et al., 1998), was not developed for this purpose, nor is it particularly advocated by its developers for this use (Knight, 1999). Ultimately, there is a great deal of overlap between disciplines (Alison et al., 2004) and typologies (Kocsis, Cooksey, & Irwin, 2002).

In the present study, only solved offenses were included, which is not optimal in prediction. It may be that certain typologies are better at predicting the previous conviction record of offenders in global populations that include undetected offenders. Although the inclusion of unreported, undetected, and unsolved cases is particularly difficult in empirical research of this nature, their exclusion should be noted as a limitation to the study because ultimately a significant proportion of sex offenders will not be included in the analyses.

Further, the use of police and prison services records as research data is not optimal. Generally, these data are not collected for empirical research and every source of crime data will be biased in one way or another (Canter et al., 2003). However, as Bennell and Canter (2002) point out, there is no obvious alternative to using such data and therefore it is generally acknowledged that crime research involving police data will have limited results (Farrington & Lambert, 1992).

In general, the study design attempted to potentially nullify the majority of possible individual biases and limitations on approaches by comparing each across the same data set. In other words, the approaches were all subject to the identical biases and limitations inherent in the current data set. Nevertheless, one approach may in reality be more influenced by the particular biases and limitations in the data set than another, and this should be noted. Nevertheless, experimentally this may be of particular interest, indicating that one approach may not be as robust or generalizable as another across certain types of data set. Ostensibly, empirical differences were found between approaches despite the data being limited, not because of it.

CONCLUSION

The current research compared Hazelwood's (1987) *Power and Anger* model, the *Behavioral Thematic* model of Canter et al. (2003), and the *MTC:R3* (Knight & Prentky, 1990) typology with a multivariate approach adapted from the thematic model of Canter et al. (2003). Logistic regression and ROC analyses revealed that both the Power and Anger and MTC:R3 models were similarly effective, while the Behavioral Thematic model was found less effective in predicting the previous convictions of these 85 stranger rapists. However, multivariate behavioral analyses far exceeded all three models. The results suggest that emphasis should be placed on further exploration at the individual variable level (e.g. bivariate relationships between crime scene elements and offender characteristics) to identify those behavioral elements within each typology, thematic representation, and multivariate approach that are most pragmatically useful. Further, a simplified bivariate approach may help to identify which variables are more influenced or moderated by other situational or contextual elements of sexual offenses.

REFERENCES

- Ainsworth, P. B. (2001). *Offender profiling and crime analysis*. Devon, UK: Willan.
- Alison, L. (2005). *The forensic psychologist's casebook: Psychological profiling and criminal investigation*. Devon, UK: Willan.
- Alison, L., Bennell, C., Mokros, A., & Ormerod, D. (2002). The personality paradox in offender profiling: A theoretical review of the processes involved in deriving background characteristics from crime scene actions. *Psychology, Public Policy and Law*, 8, 115–135.
- Alison, L., & Canter, D. (1999). Professional, legal and ethical issues in offender profiling. In D. Canter, & L. Alison (Eds.), *Profiling in policy and practice* (pp. 21–54). Aldershot: Ashgate.
- Alison, L., Smith, M. D., Eastman, O., & Rainbow, L. (2003). Toulmin's philosophy of argument and its relevance to offender profiling. *Psychology Crime and Law*, 9, 173–184.
- Alison, L., & Stein, K. (2001). Vicious circles: Accounts of stranger sexual assault reflect abusive variants of conventional interactions. *The Journal of Forensic Psychiatry*, 12, 515–538.
- Alison, L., West, A., & Goodwill, A. M. (2004). The academic and the practitioner: Pragmatists' views of offender profiling. *Psychology, Public Policy and Law*, 10, 71–101.
- Ault, R. L., & Reese, J. T. (1980). A psychological assessment of crime profiling. *FBI Law Enforcement Bulletin*, 49, 22–25.
- Beech, A., Oliver, C., Fisher, D., & Beckett, R. C. (2006). *STEP 4: An evaluation of the U.K. Prison Sex Offender Treatment Programme as used for the treatment of rapists and sexual murderers*. Birmingham: University of Birmingham. Retrieved January 20, 2008, from www.hmprisonservice.gov.uk/assets/documents/100013DBStep_4_SOTP_report_2005.pdf
- Bennell, C., & Canter, D. V. (2002). Linking commercial burglaries by modus operandi: Tests using regression and ROC analysis. *Science and Justice*, 42, 153–164.
- Bennell, C., & Jones, N. J. (2005). Between a ROC and a hard place: A method for linking serial burglaries using an offender's modus operandi. *Journal of Investigative Psychology and Offender Profiling*, 2, 23–41.
- Boon, J. (1997). Contribution of personality theories to psychological profiling. In J. L. Jackson, & D. A. Bekerian (Eds.), *Offender profiling: Theory, research and practice* (Vol. 43–59). Chichester: Wiley.
- Canter, D. V. (1994). *Criminal shadows*. London: Harper Collins.
- Canter, D. V. (2000). Offender profiling and criminal differentiation. *Legal and Criminological Psychology*, 5, 23–46.
- Canter, D. V., Bennell, C., Alison, L. J., & Reddy, S. (2003). Differentiating sex offenses: A behaviorally based thematic classification of stranger rapes. *Behavioral Sciences and the Law*, 21, 157–174.
- Canter, D., & Heritage, R. (1990). A multivariate model of sexual offence behaviour: Developments in 'offender profiling'. *The Journal of Forensic Psychiatry*, 1, 185–212.
- Copson, G., Babcock, R., Boon, J., & Britton, P. (1997). Articulating a systematic approach to clinical crime profiling. *Criminal Behavior and Mental Health*, 7, 13–17.
- Daubert v. Merrell Dow Pharmaceuticals, Inc. 509 U.S. 579, 113 S. Ct. 2786 (1993).
- DeLong, E. R., DeLong, D. M., & Clarke-Pearson, D. L. (1988). Comparing the areas under two or more correlated receiver operating characteristic curves: A nonparametric approach. *Biometrics*, 44, 837–845.

- Dorfman, D. D., & Alf, E. (1969). Maximum likelihood estimation of parameters of signal detection theory and determination of confidence intervals—rating-method data. *Journal of Mathematical Psychology*, 6, 487–496.
- Douglas, J. E., Burgess, A. W., Burgess, A. G., & Ressler, R. K. (1992). *Crime classification manual*. San Francisco, CA: Jossey-Bass.
- Farrington, D. P., & Lambert, S. (1992). *The feasibility of a statistical approach to offender profiling: Burglary and violence in Nottinghamshire*. London: Home Office.
- Fishman, D. B. (2000). The case for pragmatic psychology. *Contemporary Psychology*, 45, 108–111.
- Fleiss, J. L. (1981). *Statistical methods for rates and proportions*. New York: Wiley.
- Goodwill, A. M., & Alison, L. J. (2007). When is profiling possible? Offence planning and aggression as moderators in predicting offender age from victim age in stranger rape. *Behavioral Sciences and the Law*, 25, 823–840.
- Groth, A. N., & Birnbaum, H. J. (1979). *Men who rape: The psychology of the offender*. New York: Plenum.
- Groth, A. N., Burgess, A. W., & Holmstrom, L. L. (1977). Rape: Power, anger and sexuality. *American Journal of Psychiatry*, 134, 1239–1243.
- Hakkanen, H., Lindlof, P., & Santtila, P. (2004). Crime scene actions and offender characteristics in a sample of Finnish stranger rapes. *Journal of Investigative Psychology and Offender Profiling*, 1, 17–32.
- Hanley, J. A., & McNeil, B. J. (1983). A method of comparing the areas under receiver operating characteristic curves derived from the same cases. *Radiology*, 148, 839–843.
- Hazelwood, R. R. (1987). Analyzing the rape and profiling the offender. In R. R. Hazelwood, & A. W. Burgess (Eds.), *Practical aspects of rape investigation: A multidisciplinary approach* (pp. 169–199). New York: Elsevier.
- Hazelwood, R. R. & Burgess A. W. (Eds.), (1987). *Practical aspect of rape investigation: A multidisciplinary approach*. New York: Elsevier.
- Hazelwood, R. R., & Douglas, J. E. (1980). The lust murderer. *FBI Law Enforcement Bulletin*, 1–5.
- Hazelwood, R. R., & Warren, J. I. (1989). The serial rapist: His characteristics and victims. *FBI Law Enforcement Bulletin*, 58, 18–25.
- Hazelwood, R. R., & Warren, J. I. (2000). The sexually violent offender: Impulsive or ritualistic? *Aggression and Violent Behavior*, 5, 267–279.
- Her Majesty's Crown Prosecution Services Inspectorate (HMcpis). (2007). *Without consent: A report on the joint review of the investigation and prosecution of rape offences*. Retrieved April 18, 2007, from <http://www.inspectorates.homeoffice.gov.uk/hmic>
- Hicks, S. J., & Sales, B. D. (2006). *Criminal profiling: Developing an effective science and practice*. Washington, DC: APA.
- Homant, R. J., & Kennedy, D. B. (1998). Psychological aspects of crime scene profiling: Validity research. *Criminal Justice and Behavior*, 25, 319–343.
- Klassen, D., & O'Connor, W. A. (1994). Demographic and case history variables in risk assessment. In J. Monahan, & H. J. Steadman (Eds.), *Violence and mental disorder: Developments in risk assessment* (pp. 229–259). Chicago, IL: University of Chicago Press.
- Knight, R. A. (1999). Validation of a typology for rapists. *Journal of Interpersonal Violence*, 14(3), 303–330.
- Knight, R. A., & Prentky, R. A. (1990). Classifying sexual offenders: The development and corroboration of taxonomic models. In W. L. Marshall, D. R. Laws, & H. E. Barbaree (Eds.), *Handbook of sexual assault: Issues, theories, and treatment of the offender* (pp. 23–52). New York: Plenum.
- Knight, R. A., Warren, J. I., Reboussin, R., & Soley, B. J. (1998). Predicting rapist type from crime-scene variables. *Criminal Justice and Behavior*, 25, 46–80.
- Kocsis, R. N., Cooksey, R. W., & Irwin, H. J. (2002). Psychological profiling of offender characteristics from crime behaviors in serial rape offenses. *International Journal of Offender Therapy and Comparative Criminology*, 46, 144–169.
- McGrath, M. G. (2000). Criminal profiling: Is there a role for the forensic psychiatrist? *Journal of the American Academy of Psychiatry and the Law*, 28, 315–324.
- MedCalc Software. (2008). *MedCalc for Windows*, Version 9.2.1.0. Mariakerke, Belgium.
- Muller, D. A. (2000). Criminal profiling: Real science or just wishful thinking? *Homicide Studies*, 4(3), 234–264.
- National Centre for the Analysis of Violent Crime. (1990). *Criminal investigative analysis: Sexual homicide*. Quantico, VA: NCA VC FBI Academy.
- Prentky, R. A., Cohen, M., & Seghorn, T. K. (1985). Development of a national taxonomy for the classification of rapists: The Massachusetts Treatment Center system. *American Academy of Psychiatry and Law*, 13, 39–70.
- SPSS Inc. (2006). *SPSS for Windows*, Version 14.0. Chicago, IL.
- Swets, J. A. (1988). Measuring the accuracy of diagnostic systems. *Science*, 240, 1285–1293.
- Wilson, P., Lincoln, R., & Kocsis, R. (1997). Validity, utility and ethics of profiling for serial violent and sexual offenders. *Psychiatry, Psychology and Law*, 4, 1–12.

Copyright of Behavioral Sciences & the Law is the property of John Wiley & Sons, Ltd. 2004 and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.