1997

The Predictive Accuracy of the Violent Offender Treatment Program Risk Assessment Scale

Ann Ward
Edith Cowan University

Edith Cowan University

Copyright Warning

Running Head: Assessment of the VOTP Risk Screening Device

The Predictive Accuracy of the Violent Offender Treatment Program Risk Assessment Scale

Ann Ward

Edith Cowan University

A Thesis Submitted in Partial Fulfilment of the

Requirements for the Award of

Bachelor of Arts (Psychology) Honours

Faculty of Health and Human Sciences, Edith Cowan University.

USE OF THESIS

The Use of Thesis statement is not included in this version of the thesis.

Abstract

Current methods for screening violent offenders for program eligibility are expensive and time consuming. Developers of the Violent Offender Treatment Program (VOTP) have designed a brief and economical instrument to screen offenders for program eligibility. The present study was undertaken to assess the reliability and predictive accuracy of the VOTP Risk Assessment Scale (RAS). An inter-rater reliability of 20 court histories attained a mean kappa of .81. The RAS was applied to court histories of 202 violent offenders released between 1985 and 1987. A 10-year follow-up of convictions for violent behaviour yielded a 47% base rate. Receiver Operating Characteristic (ROC) curves showed that for varying time-at-risk periods the predictive accuracy remained between .72 and .76. The recommended cutoff score for all time-at-risk periods was 11. The relatively high accuracy rate of the VOTP RAS indicated that it was accurate enough to aid program eligibility decisions.

Declaration

I certify that this thesis does not incorporate, without acknowledgment, any material previously submitted for a degree or diploma in any institution of higher education and that, to the best of my knowledge and belief, it does not contain any material previously published or written by another person except where due reference is made in the text.

Signature: _____

Date: 9-3-1998

Acknowledgments

I would like to thank the following people:

Steve Baldwin, What can I say? Thank you. Your advice, support, humour, honesty, and guidance were much appreciated. Your intuition which undoubtedly guided your criticisms served us both well. Remember that car driving down the highway with me in it, Led Zeppelin on 10? Well I have news for you, You're driving! Oh, and sorry, but I think you should leave Orb at home.

I would also like to thank my advisors at Edith Cowan University, Guy Hall and Julie Thacker for their ideas, advice and no doubt numerous unwanted telephone calls.

Thanks to Graham Chapman at the Alternatives to Violence Unit at the Ministry of Justice. Your support and guidance did not go unnoticed. Thanks to the others for their help in retrieving records, getting me into the computers and what not. Thanks Marlene, Janet, Trish and others.

An *extra* special thank you to Dr. Douglas Mossman of Wright State University in Dayton, Ohio. Thank you for the time and knowledge you gave me. Unfortunately not all of it sank in. Thank you for answering and returning all the telephone calls. And thank you for putting up with all the stupid questions.

Thanks to those at Casuarina Prison for their participation and cooperation. This thanks includes Bruce Watt and John Dockerill regardless of whether they are at Casuarina, the Ministry of Justice, Uni, or wherever.

Thanks to my peers for their undaunting support, and ears for bitchin' into. Without you guys I probably would not have made it.

And last but by no means least, thanks to Paul for putting up with the grumpiness, the one-sided conversations, the ear-bashings, and the weirdness "Tell me it's a trick Dr. Giggle". We both have the same understanding on that one. None!

TABLE OF CONTENTS

List of Tables

List of Figures

Introduction

Background

Risk assessment is a screening device used to classify offenders (Champion, 1994). Offenders are often screened several times as they move through the justice system. For example, a risk assessment instrument may be used to help sentencing decisions (Andrews & Bonta, 1994). If an offender is classified as high-risk (i.e., they are likely to reoffend) then they may receive a more severe sentence than an offender who is deemed low-risk. Likewise, risk assessment is used to aid decisions about the level of supervision in prison and in the community (Baird, 1981), program eligibility (Brown, 1996), and parole (Baird, 1981).

The many uses of risk assessment have necessitated a myriad of tools. For example, certain risk assessment instruments may be used for predicting probation outcome (Harris, 1994), parole release (Hoffman, 1983), inmate classification (Loza & Simourd, 1994) and violence (Klassen & O'Connor, 1988a). Furthermore, different countries and states adopt their own risk assessment tools (see Champion, 1994).

Risk assessment for violent offenders is necessary for several reasons. The continual perpetration of violent offences by individuals undermines public safety (Shah, 1981), increases prison overcrowding (Bonta & Motiuk, 1990) and drains fiscal resources (Morgan, 1994). As a result, it has become necessary to intervene in the cycle of violence committed by high-risk offenders (i.e., those who are likely to commit further violent offences). However, difficulties arise when attempting to

reliably identify high-risk violent offenders. The identification of high-risk violent offenders is necessary however, to divert them into treatment programs.

The designation of violent offenders to treatment programs is of utmost importance. Not only are there are a limited number of programs for violent offenders (Ward & Baldwin, in press), but there are a limited number of places available in those programs. As a result it becomes necessary to develop criteria to discriminate between offenders who are likely to commit future violent offences and those who are not. These criteria comprise risk assessment scales for violent offenders.

Many other instruments assess the risk or risk/needs of offenders (e.g., Bonta & Motiuk, 1985; Harris, Rice & Quinsey, 1993; Klassen & O'Connor, 1988a, 1988b). Such instruments have been developed specifically for the intended population (i.e., psychiatric, general recidivists). However, to date there is no instrument to discriminate reliably between low-risk violent offenders and high-risk violent offenders. As it is simply not viable to assess all offenders, it is necessary to discriminate between low- and high-risk violent offenders to aid the decision of whether to assess for program eligibility.

Theoretical Overview

Defining Risk, Dangerousness, and Violence

One of the main aims of psychologists is to predict human behaviour. This has become particularly important in the forensic field. Legal systems have required that psychologists and psychiatrists predict dangerousness, risk, and violence. The

prediction of such behaviours has caused much controversy (Grisso & Appelbaum, 1992). Similarly, definitions of the terms 'dangerousness', 'risk', and 'violence' have caused confusion.

Shah (1981) defined dangerousness as the propensity to engage in "...acts that are characterized by the application or overt threat of force and are likely to result in injury to other persons" (p. 152). Shah (1981) stated that dangerous behaviours and violence were "synonymous" (p.152). Champion (1994) however, defined dangerousness as "propensities to cause harm to others and oneself" (p. 19). He further stated that the terms dangerousness and risk were "interchangeable". The integration of these two definitions has created an anomalous state in which dangerousness, risk, and violence can be construed as the same thing. Furthermore, the concept was extended not only to include others but also oneself. However, the issue of 'self harm' is a large and complex area and will not be dealt with here.

Prins (1996) made an important distinction between the concepts of risk and dangerousness. He stated that risk should be viewed as the "likelihood" that an event would occur. Whereas, dangerousness referred to the "degree of damage (harm) that may occur should the event happen" (p.45). This distinction between risk and dangerousness is vital. Viewing risk and dangerousness as the same thing, suggests that if an event were to occur, then it would be a harmful event.

Consolidating on Prins' (1996) definition, if the degree of harm constitutes dangerousness, then someone who is violent is not necessarily considered dangerous. For example, common assault is deemed to be an act of violence, however it does not result in bodily harm. Thus dangerous behaviour is violent behaviour, but violent behaviour is not necessarily considered to be dangerous behaviour. This rationale

shows that violent behaviour covers a broader spectrum of behaviours than the concept of dangerousness. Therefore, risk refers to the likelihood of future offences, dangerousness to the severity of the offence, and violence to overt threat or force.

Measuring Violence

Both Australian and overseas studies (Broadhurst & Maller, 1990; Kraus, 1979; Selby, 1984; Shaffer, Waters, & Adams, 1994) have offered parameters governing violence. The purpose of a study often determines the definition of violence. Whilst some studies only include physically harmful violent behaviour (e.g., Kraus, 1979), others include threats of violent behaviour (e.g., Shaffer et al., 1994).

In an Australian study, Kraus (1979) classified murder, manslaughter, major assault (assaults requiring hospital treatment), rape, robbery and arson as violent crimes. However, Broadhurst and Maller (1990) in their recidivism study, listed assaults receiving sentences of less than six months in the top six most severe crimes. Overseas researchers such as Selby (1984) and Shaffer et al., (1994) have extended violent crimes to include verbal threats. Although there appears to be agreement in cases of extreme violence (i.e., murder, rape), common assaults and verbal threats are, through omission, on occasions considered as non-violent.

Defining violence by physical contact excludes armed robbery as a violent offence. However, armed robbery is considered to be a serious and violent offence. A study by Indermaur (1990) showed that armed robbery was rated by both the community and judges as more severe than assault and grievous bodily harm. Some

researchers (e.g., Hamparian, Schuster, Dinitz, & Conrad, 1978) have defined violence by criminal legislation.

In the present study, violence is considered to include both acts of violence (i.e., crimes ranging from murder to common assault) and fear-inducing behaviours (i.e., threats). Any attempts at such crimes are also considered to be violent. As the crime of arson is viewed as a property offence, it has been omitted.

Actuarial Vs Clinical Predictions

Traditionally, risk of violence (usually among mentally disordered patients) was subjectively judged by the attending clinician (i.e., psychologist or psychiatrist). However, in his seminal work, Monahan (1981) found that clinicians' predictions were highly overrated. Clinical predictions of violence "seemed to be wrong at least twice as often as they were right" (Monahan, 1984, p.10). This in part was due to the lack of guidelines for clinicians to follow in order to predict violence (Monahan & Steadman, 1994).

Since clinicians had such a poor record of predicting violence, Monahan (1984) called for "second-generation thinking" in the prediction of violent behaviour. Second-generation thinking was typified by: (a) the belief that violent behaviour could be predicted, (b) the question of how accurately violent behaviour could be predicted, and (c) that prediction be viewed in relative (rather than absolute) value. Monahan (1984) suggested that perhaps violent behaviour could be predicted "accurately enough to be useful in some policy decisions" (p.11).

Monahan (1984) also called for the use of actuarial methods. Actuarial methods are based on group characteristics (Champion, 1994) that are statistically

derived. For example, based on actuarial methods, relevant characteristics of a parolee may be compared to what has been found to be characteristic of other similar parolees. Likewise, a probationer's characteristics would be compared to other probationers'.

Monahan's (1984) call for second-generational thinking was answered by a flood of research based on actuarial methods (e.g., Bonta & Motiuk, 1992; Gardner, Lidz, Mulvey & Shaw, 1996a; Gardner et al., 1996b; Klassen & O'Connor, 1988a, 1988b; Rice & Harris, 1992, 1995a, 1995b; Webster, Harris, Rice, Cormier, & Quinsey, 1994). Such studies have provided relatively promising results.

However, with the development of guidelines for clinicians (Borum, 1996) and methodological advances (Monahan & Steadman, 1994), clinical predictions have once again been pitted against actuarial predictions. Gardner et al., (1996a) found that even with such advances, actuarial predictions based solely on the individual's history of violence were more accurate than clinical predictions. Even though actuarial predictions have been found to be more accurate than clinical predictions, results remain at best, modest.

Public Policy and the Risk Principle

Public policy influences risk assessment in two ways. It influences both the development and the subsequent use of risk assessment scales (Monahan, 1981, 1984, 1992; Monahan & Steadman, 1994). The development of a risk assessment scale is dictated by the type of behaviours it is attempting to predict. Hence it is necessary to determine which behaviours are included, and to what extent such

behaviours are unacceptable to society. Thus initiation of policy is necessary in order to guide the task of prediction.

Historically, in the late 19th century, dangerousness was a concept associated with habitual offenders (Pratt, 1995). The type or severity of offence that the habitual offender committed was irrelevant. Thus many petty thieves were considered dangerous simply because they were persistent offenders. However, during the 1950s, the sexual psychopath laws were passed in the United States of America (USA) (Pratt, 1995). The 1950s sexual psychopath laws were followed by provisions for 'dangerous offenders'. Such provisions often included indeterminate sentencing. During the 1970s, most states of the USA practiced indeterminate prison sentencing (Monahan, 1984). However, most states have since abolished the practice (Monahan, 1984).

Throughout the 1970s, many countries introduced or refined their dangerousness-related laws to pertain almost exclusively to violent/sexual offending (Pratt, 1995). Canadian 'dangerousness' laws were targeted at high-risk, repeat violent offenders. Similarly, New Zealand and Victorian legislation was targeted at the serious sexual and serious violent offender (Pratt, 1995). However, in Western Australia (WA), laws pertaining to dangerousness remained broad and largely ill-defined. For example, detention at His Governor's Pleasure is still in use. This indeterminate sentence was specifically designed to deal with the habitual criminal rather than the dangerous offender. Although rarely used, it has on occasion been used for the dangerous offender.

Shah (1981) acknowledged that the dangerousness concept was based in public policy. Such policies were based upon the Uniform Crime Reports developed

by the Federal Bureau of Investigation (FBI). The FBI categorized crimes of violence as: murder, aggravated assault, rape, robbery and any attempts of such crimes (Shah, 1981). This categorization leads to the conclusion that criminal offences such as common assault and threats to murder, are not acts of violence.

However more recently, the FBI's National Centre for the Analysis of Violent Crime, published the Crime Classification Manual (Ressler, Douglas, Burgess & Burgess, 1992). The classification of violent crimes included homicide, arson, and rape and sexual assault. This manual was designed specifically for the classification of such crimes (e.g., domestic homicide, contract killing, kidnap murder). By omission, crimes such as robbery and assaults, were no longer classified as violent. Furthermore, there appears to be no justification for the inclusion of arson, which has traditionally been considered a property offence. However, this may once again be related to current public policy regarding dangerousness.

Exactly what constitutes a 'dangerous' (as opposed to habitual) offender in WA remains undefined. However, a legislative act has been introduced to deal with serious violent offenders. Section 37A(a)(b) of the Imprisonment and Parole Act (WA) alludes to the concept of dangerousness. People considered dangerous were offenders who had been sentenced to prison for five years or more for crimes against the person (Broadhurst, 1991). However, being imprisoned for five or more years for crimes against the person did not automatically categorize an individual as dangerous. Such categorizations were left to the discretion of the parole board (Broadhurst, 1991). Legislation such as this leaves each case to be considered individually.

The use of a risk assessment scale may also be affected by policy, by dictating the cutoff point on the scale (Monahan, 1992; Mossman & Somoza, 1992). For example, if current policy suggested that probation be used as often as possible in an effort to reduce prison overcrowding and ease the fiscal burden of imprisonment, the cutoff point on a scale predicting probation outcome may be raised. By raising the cutoff point, more lower-risk offenders would receive probation instead of imprisonment. Similarly, if a 'get tough' policy were introduced, the cutoff point may be lowered so that more high-risk offenders receive terms of imprisonment rather than probation.

Due to the lack of explicit public policy in WA, it remains unclear as to which criminal behaviours should be targeted for intervention. The intuitive stance would be to target all criminal behaviours, violent or otherwise. However, such a stance is unfeasible simply because fiscal resources are limited. With such restrictions in mind, one of the first subgroups of offenders to warrant intervention would be violent offenders. During the past 10 years violent offenders in WA have *en masse* received a programme package (Skills Training and Aggression Control) (Howells, Watt, Hall & Baldwin, 1997). Such treatment strategies may not only result in a fiscal drain, but may also in certain cases be ineffective. Howells (1996, p.73) stated "there is a distinctly farcical quality to observing a "psychopathic" offender, who may have used violence instrumentally and coldly throughout his life, being persuaded to engage in relaxation exercises in order to control his temper!" Such treatment may not simply be ineffective, but according to the 'risk principle' (Andrews, Kiessling, Robinson, & Mickus, 1986; Andrews, Robinson, & Balla, 1986; Andrews & Bonta, 1994) it may prove to be detrimental.

The underlying assumption of the risk principle is that criminal behaviour can be predicted. It is based on the idea that level of treatment services should be matched to the risk level of the offender (Andrews & Bonta, 1994). For example, higher-risk offenders should receive intensive intervention whereas low-risk offenders should receive minimal (if any) intervention (Andrews, Kiessling et al., 1986).

In a series of studies by Andrews and colleagues (Andrews, Kiessling et al., 1986: Andrews, Robinson, & Balla, 1986; Andrews, Bonta & Hoge, 1990; Andrews, Zinger, Hoge, Bonta, Gendreau, & Cullen, 1990), a reduction in recidivism for high-risk offenders who had undergone intensive intervention was found. Conversely, intensive intervention for low risk offenders had either a minimal or detrimental effect on recidivism.

The risk principle has to date, only been applied to general recidivists whose records were free of violence. However, until the risk principle has been excluded as being applicable to violent offenders, it must be considered.

Risk and General Recidivism

General recidivism refers to the commission of subsequent crimes by offenders. There is a vast literature on general recidivism, much of which has focused on delinquency. General recidivism covers offences ranging from drink-driving to wilful murder. The focus of general recidivism studies may include: prevalence of offence types (i.e., sex offences, drug offences) (Broadhurst & Maller, 1990), and offender characteristics in relation to sentence (Broadhurst & Maller, 1990; Walker, 1989). Such studies have shown a number of correlates between crime and offender

characteristics such as: age, gender, race, socio-economic status, and educational level. Subsequently, such variables have been incorporated into many risk assessment scales as predictors.

In a recent study, Katsiyannis and Archwamety (1997) examined factors related to recidivism among delinquent youths. The sample consisted of males between 12 and 20 years of age (n = 294). The recidivists (n = 147) were compared to non-recidivists (n =147) on institutional, cognitive and academic achievement variables. Of the 23 variables significant differences between groups were: length of first stay at the facility; reading, writing, and math levels; age at first commitment; age of first offence; and risk score as estimated by the facility.

Level of Supervision Inventory

However, some risk assessment scales such as the Level of Supervision Inventory (LSI) have been developed largely ignoring such variables. In a series of studies by Bonta and colleagues (Bonta and Motiuk, 1985, 1987, 1990, 1992; Motiuk, Bonta & Andrews, 1986; Motiuk, Motiuk & Bonta, 1992), the LSI was shown to be useful in predicting recidivism, institutional performance and inmate classification.

The LSI is a risk/needs assessment scale (Bonta & Motiuk, 1987). It consists of 11 categories and 58 items. The LSI is one of the few risk assessment scales claimed to be based on theoretical premises. Bonta and Motiuk (1987) stated that the LSI was derived from social learning theory. However, this is not always recognizable in the scale. For example, it is arguable that attitudes and orientation

(e.g., supportive of crime; unfavourable towards convention; poor, toward judicial sentence; and poor, toward supervision) are learnt socially. However, the connection of financial and accommodation items to social learning theory is tenuous at best.

Theoretical derivation aside, the LSI variables appear to be better predictors for recidivism, institutional performance or inmate classification than scales such as the Megargee-Minnesota Multiphasic Personality Inventory (M-MMPI) (Motiuk et al., 1986). Motiuk et al., (1986) evaluated the predictive accuracy of the LSI against the predictive accuracy of the M-MMPI. No significant difference between high- and low-risk offenders was found when using the M-MMPI to measure halfway house outcome and reincarceration. However, significant differences were found between high- and low-risk offenders when using the LSI.

Other Applications of the LSI

The LSI has been subjected to investigations in a variety of contexts. For example, Bonta and Motiuk (1985) found that it accurately classified between 67% to 75% of offenders diverted to halfway houses. A paper-and-pencil self-report inventory (SRI) adapted from the LSI showed that the SRI correctly identified a higher proportion of offenders reincarcerated (66.7%) than the LSI (53.5%). Similarly the SRI performed better than the LSI for misconducts (SRI = 72.5%, LSI = 62.5%). Whilst the LSI had better predictive accuracy for assaults (61.2%), it was only marginal (SRI = 60.6%).

The LSI has also been examined for use in female offender populations (Coulson, Ilacqua, Nutbrown, Giulekas & Cudjoe, 1996). Coulson et al., (1996)

found that by lowering the cutoff score between high- and low-risk offenders, the LSI was a useful tool for the prediction of recidivism in female offenders.

Statistical Information on Recidivism

Research by Bonta, Harman, Hann, and Cormier (1996) is salient to the present study. In a re-validation study of the Statistical Information on Recidivism (SIR) scale, Bonta et al., (1996) employed the use of Receiver Operating Characteristic (ROC) curves to assess the predictive accuracy of the SIR scale for general recidivism, broadly-defined violent recidivism, and narrowly-defined violent recidivism.

Bonta et al., (1996) used a modified version (two items were dropped) of the SIR scale as information for some of the items were unobtainable from official records. The narrowly-defined violent recidivism category included: homicide, sexual assault, and aggravated sexual assault. The broadly-defined violent recidivism category included: offences from the narrowly-defined category, robbery, weapon offences and less serious sexual assaults. The general recidivism category included all other offences.

The SIR scale consists of 13 items. Each item has sub-items, some are binary whilst others are in continual and categorical form. The items that comprise the SIR are: current offence; age at admission; previous imprisonment; previous parole breach; escape history; maximum security classification; age at first adult conviction; previous assault; marital status; interval at risk since last offence; aggregate sentence; previous violent sex offence; and, previous break and enter.

Bonta et al., (1996) found that all SIR items could predict general recidivism, except the previous conviction for violent sex offence, which could only predict the narrowly-defined violent recidivism. Predictive accuracy for general recidivism was 74%. For broadly-defined violent recidivism the accuracy rate was somewhat lower (64%). Narrowly-defined violent recidivism was correctly predicted in 65% of cases.

The literature on general recidivism has offered some guidance for violent recidivism studies. Whilst some general recidivism scales (e.g., the SIR) have been applied to the prediction of violent recidivism, others such as the LSI remain to date untested.

Risk, Violence and the Mentally Disordered Offender

Most of the literature dealing with future violent offending has focused in the area of the mentally disordered offender. Under most circumstances, predictions of violence are not tested. However, researchers Steadman and Cocozza (1974) had the opportunity to test clinical predictions of violence using actuarial methods.

Applying actuarial methods, Steadman and Cocozza (1974) found that age and Legal Dangerousness Score (LDS) were the best predictors of future violence. The LDS consisted of four items: presence of juvenile record, number of previous arrests, presence of convictions for violent crimes, and severity of offence. Thornberry and Jacoby (1979) who conducted similar research, concluded that whilst age was the single most powerful predictor of dangerousness, it was still extremely inadequate.

In his 1981 review, Monahan, like Thornberry and Jacoby (1979), acknowledged criticisms of clinical prediction research. Criticisms suggested (a) that

studies tested factors other than violent behaviour, such as bureaucratic inertia, (b) that by the time predictions were tested they were out-dated, and (c) that a high level of violence went undetected. Whilst valid, Monahan (1981) maintained that such criticisms did not negate the value of the prediction of violent behaviour.

Risk Instruments and the Mentally Disordered Offender

In a sample of male mental health admissions, Klassen and O'Connor (1988a) examined violent recidivism. Twenty two variables were sorted into seven categories: arrest record, mental health records, demographics, family background, previous violence, test scores (abstract reasoning and Shipley-Institute of Living), and situational measures over the preceding three months. All arrest records, test scores and family background variables were statistically significant. Only one each of the mental health records (assault in presenting problem), demographics (never married) and previous violence (number of violent incidents in past year) were significant. Three of the eight situational variables were significant.

From a three- and six-month follow-up, the instrument showed a high correct classification figure of 85.3%. Of the remaining 239 participants, a total of 93.9% nonviolent offenders were correctly classified and 59.3% of violent offenders were correctly classified. Whilst there was a low rate of false negative (6.1%), there was a high rate (40%) of false positives.

Another study by Klassen and O'Connor (1988b) of 304 mental patients showed similar classification results. Mental hospital re-admission records, arrest records, and interviews were examined to determine violence in schizophrenic and non-schizophrenic inpatients. Historical (e.g., criminal record) and situational

(e.g., current needs) variables were examined to determine criteria for violence. A 1-year follow-up showed that variables related to subsequent violence for both groups included: early family history variables, employment variables, current relationships, and hospital record items (e.g., violence in presenting problem).

Variables that demonstrated consistent differences between the schizophrenic group and the nonschizophrenic group were: life events items, and substance abuse (positive associations for non-schizophrenics, negative associations for schizophrenics). Schizophrenics were correctly classified in 88.1% of cases, with a false positive rate of 17.6%. Non-schizophrenics showed a 92.9% correct classification, with a lower false positive rate of 9.4%. It was also found that arrests for violence were only predictive for the non-schizophrenics.

Correct classification results for both Klassen and O'Connor studies (1988a; 1988b) were similar, there was a notable decline in the false positive rate. However, whether the decline in false positive rate was an artifact of sample segregation (i.e., schizophrenic/non-schizophrenic) variable selection, or other factors, is unknown.

In a study of psychopathy and violent recidivism, Harris, Rice and Cormier (1991) used institutional files for data collection. From an extensive pool, variables were categorized under: childhood history, adult adjustment, offense and assessment, and institution and program. Over the 5% of records rated (n = 10), retained continuous variables attained a mean interrater agreement of $\kappa = .83$. Over a 10-year (M = 124.5 month) follow-up period, 169 of the 176 maximum security hospital inmates had an opportunity to reoffend. Of those that did fail (i.e., reoffend), the mean time-at-risk until failure was 55.4 months.

Harris et al., (1991) found that those diagnosed as psychopaths had a higher rate of violent recidivism than non-psychopaths. However, violent recidivism differences between groups could not be attributed to past criminal behaviour alone. Harris et al., (1991) did find that the Psychopathy Checklist had predictive utility for violent recidivism, even though most of the sample had past violent offences.

Similar to Klassen and O'Connor's (1988b) study of schizophrenics and nonschizophrenics, Rice and Harris (1992) examined general, and violent recidivism (n = 96 matched pairs) in relation to schizophrenia. Variables similar to the study by Harris et al., (1991) were utilized. Over a 7-year follow-up, Rice and Harris (1992) found that schizophrenics had a significantly lower rate of general recidivism (35%) than the non-schizophrenics (53%). As with Klassen and O'Conner's (1988b) study, a significant difference for violent recidivism between groups was not found. Most schizophrenic recidivism failures (44%) were found to occur whilst the participant was a patient in an open psychiatric hospital.

Rice and Harris (1992) found that the best predictors were previous offence history, age, marital status, alcohol abuse, and past aggressive behaviour. Two of the variables (the LSI and the Psychopathy Checklist) predicted both violent and general recidivism amongst schizophrenic patients.

A statistical instrument developed by Harris, Rice and Quinsey (1993) for mentally disordered offenders, included four areas: childhood history, adult adjustment, index offence, and assessment. A total of 42 variables comprising the four areas were examined. Of the 42 variables only 12 were included in the final scale: Psychopathy Checklist, separation from parents under the age of 16, victim injury at index offence, *DSM-III* schizophrenia, never married, elementary school

maladjustment, female victim-index offence, failure on prior conditional release, property offence history, age at index offence, alcohol abuse history, and *DSM-III* personality disorder.

The instrument developed by Harris et al., (1993) showed a classification accuracy of approximately 75%. Of the 618 patient sample, 67 men were detained. The biggest difference between those released and those detained was the severity of index offence. However, the seriousness of index offence variable was negatively correlated with violent recidivism. Similarly, injury to the victim, and having a female victim, were also negatively correlated with violent recidivism. Detainees however, were more likely to have had both a female victim and to have injured their victim, than the released group. The only indication that those detained would have a higher rate of violent recidivism according to the instrument, was that fewer of the detained patients had been married.

Harris and Rice (1995) later reanalyzed their 1993 data using Receiver Operating Characteristic curves. Correct classification for broadly-defined violence (physical attacks and fear-inducing behaviour) ranged between 68% and 70% at varying time-at-risk periods (i.e., 3.5, 6, and 10 years). Correct classification for narrowly-defined violence (physical attacks with a record of more than one common assault) ranged between 71% and 81%. The Area Under the Curve (AUC) was .76.

McNiel and Binder (1994) also studied inpatient violence. The Screening Checklist, a brief instrument to be used at admission, consisted of five items. The first item assessed physical attacks or fear-inducing behaviour two weeks prior to admission. The second item assessed the absence of suicidal behaviour, also two weeks prior to admission. The third item related to diagnosis (e.g., schizophrenia or

manic), the fourth item to gender, and the fifth to marital status. Items were structured so that a positive answer would be indicative of violence.

A ROC curve was used to help identify the optimal cutoff score. A score of two or less indicated low-risk, whilst a score of three or more indicated high-risk. Of the 238 calibration sample, 120 patients were deemed low-risk and 118 patients were deemed high-risk. The validation sample of 338 showed 201 low-risk and 137 high-risk patients. The Screening Checklist showed a total predictive value of 65.4%, with a false positive rate of 42.3%. When limited to physical attacks, the Screening Checklist had a false positive rate of 67.9%, and a total predictive value of 61.8%.

In a study that examined violence in a mental hospital sample and a prison sample, Schaffer, Waters and Adams (1994) examined variables that would discriminate violent offenders from nonviolent offenders. The variables included: age, race, marital status, vocational stability, education, socioeconomic status, developmental family, juvenile arrest history, adult arrests, mental health history, intelligence, MMPI scales, level of arousal, and violence history.

The best single predictor that appeared for both groups was the psychiatric hospital history. Other variables for the prison sample were MMPI Scales F (distress) and 1 (somatic complaints), juvenile arrest history, violence history and marital status. For the hospital sample, best predictors for violence were race, MMPI Scales 0 (socialization) and 6 (paranoia), and vocational stability. For the total sample, overall best predictor variables were: Psychiatric hospital history, race, age, vocational stability, marital status, juvenile arrest history, and MMPI Scales 0 and 6.

The overall accuracy rate for the total sample was 75%, with a false positive rate of 26%. Correct overall classification for the prison sample was 78% with a false positive rate of 20%. The overall correct classification for the hospital sample was 74%, with a false positive rate of 24%.

Accuracy of risk studies on mentally disordered offenders since the mid-1980s has ranged from 61.8% overall accuracy (McNiel and Binder, 1994) to 85% overall accuracy (Klassen and O'Connor, 1988a). Overall false positive rates ranged from 67.9% (McNiel & Binder, 1994) to 26% (Schaffer et al., 1994). Since the mid 1980s there has been an accuracy improvement on the lower end of the scale. The study by McNiel and Binder (1994) showed the lowest overall accuracy rate, and the highest false positive rate. This may be a function of either short-term prediction or the brief nature of the instrument. However, in emergency situations, it may simply be impossible to begin administering lengthy batteries of tests.

Risk and Violent Offenders

Research about mentally disordered offenders and general recidivism has extended to research on violent recidivism. However, the literature dealing with the risk of violent recidivism is limited.

One of the earlier studies of violence among an offender population was a study by Wenk, Robison and Smith (1972). Using a sample of 4,146 California Youth Authority (CYA) wards, Wenk et al. (1972) found that only 250 of the sample were incarcerated for a violent offence. Using variables such as: case histories, diagnosed clinical conditions, IQ, maturity level, grade achievement, MMPI, and the

California Personality Inventory (CPI), Wenk et al. (1972) attempted to predict future violence.

During a 15-month follow-up after release, 104 offenders were reincarcerated for a violent offence. It was found that offenders with a more than 'mild' opiate usage had a higher rate of violence than the general offender population. Similarly, offenders who had previous CYA commitments also had a higher rate of violence. However, neither group had a higher than average number of offenders committed for violent offences. Multiple offenders were found to have a higher than average rate of violence. For the entire population, only 7.4% of recidivism was violent.

In one of the most comprehensive studies of violence specialization, Brennan, Mednick and John (1989) examined a Danish cohort of 28,884 men. Using arrest records from 1944, Brennan et al., (1989) compared the specialization of violent offending with property offending up to 1974. Of the cohort, only 5,854 were retained as all other cohort members had committed either no offences, traffic offences, or special law offences (e.g., keeping shops open after certain hours). Of the remaining cohort members, 735 were found to have committed at least one violent offence and 147 had committed two or more offences.

In an effort to overcome earlier problems due to the way specialization was defined, Brennan et al., (1989) examined three aspects of violent offending: probabilistic (the probability of a first-time offenders' subsequent offence being violent), sequential (the probability of a previously classified violent offender committing a violent offence), and distributional violence (the probability that multiple violent offenders would commit a further violent offence).

Resulting evidence suggested that specialization for violent offending did exist. Whilst the probabilistic aspect provided little to suggest specialization, the sequential and distributional aspects did indicate violence specialization. Brennan et al., (1989) concluded that the results were not purely an artifact of an increased number of arrests.

Virkkunen, DeJong, Bartko, Goodwin, and Linnoila (1989) examined psychobiological variables in relation to violent recidivism. Virkkunen et al., (1989) conducted a three-year follow-up of 58 violent offenders and impulsive fire-setters. It was found that violent offenders and arsonists that committed new offences had significantly lower concentrates of 5-hydroxyindoleacetic acid, homovanillic acid, and blood glucose nadirs following glucose challenge, than nonrecidivists. Discriminant analysis yielded an 84% correct classification of violent recidivists and arsonists based on the blood glucose nadirs and 5-hydroxyindoleacetic acid.

A follow-up study of 348 males convicted of attempted manslaughter, manslaughter and arson was conducted by DeJong, Virkkunen and Linnoila (1992). Variables included the *DSM-III* axis I and axis II diagnostics, parental alcohol abuse, attempted suicide, suicide of relative, and impulsive index crime.

The single most powerful predictor for the manslaughter/attempted manslaughter group was impulsivity. Of the offenders who committed an impulsive index crime, 37.8% were recidivists. In contrast, only 5.7% of offenders whose index offences were not impulsive, committed a further crime. When used in conjunction with other variables (e.g., age, IQ) sensitivity was poorer than using impulsivity alone (only 12 of the 59 violent recidivists were identified). DeJong et al., (1992) stated

that using impulsivity as the sole predictor would have yielded a sensitivity of 90%, a false positive rate of 62.2%, and a false negative rate of 5.7%.

Amongst the arsonists, the single most powerful predictor was attempted suicide. Nearly half (15 of 34) of those that attempted suicide reoffended violently. The use of suicide as the sole predictor would have yielded a sensitivity of 68.1%, a false positive rate of 55.9%, and a false negative rate of 10.6%.

In a more recent study, Martinez (1997) examined official records of 322 offenders. Variables included: race, age in 1973 (cohort group), neighbourhood socioeconomic status, age at first arrest, violent juvenile felony conviction, prior adult felony conviction, prior adult felony conviction, alias, arrest-free (refers to a 5 year period before 1973), and incarceration period. The original 17-year follow-up period was reduced to 14 years after mean time served in prison was subtracted.

Age of first arrest, the use of an alias, and prior incarceration were found to be the best predictors of criminal activity. For violent crimes the best predictors were race, prior adult felony, and the use of an alias. Of the entire cohort, only 25% committed further violent offences.

Studies of violence among an offender population have in a number of cases (e.g., DeJong et al., 1992; Wenk et al., 1972) used clinical diagnosis as a predictor. DeJong et al., (1992) found impulsivity to be the single most powerful predictor for future violence. However, in most studies among an offender population, as with a mentally disordered population, static variables have been found to be equally as good (if not better) predictors than dynamic variables.

Static Vs Dynamic Variables

Andrews and Bonta (1994) identified two types of predictor variables: static variables and dynamic variables. Static variables are those that cannot be changed (e.g., criminal history). Dynamic variables are changeable (e.g., companions, employment). Whilst the literature reflects little disagreement regarding the value of certain static predictor variables such as: age, gender, criminal history, and early family factors, no such agreement exists in regard to dynamic variables (Gendreau, Little, & Goggin, 1996).

The utility of static variables such as past offending has been shown repeatedly. Klassen and O'Connor (1994) noted that almost any measure of past offending may be predictive of future violence. For example, variables that have been found to correlate with violent recidivism include: age of offending onset (Elliot, 1994; Webster, Harris, Rice, Cormier, & Quinsey, 1994), number of prior convictions (Gilmore & Walkey, 1981), history of escapes from institutions (Webster et al., 1994), previous arrests for violent crimes (Klassen & O'Connor, 1988a; Thornberry & Jacoby, 1979), juvenile violence (Wenk, Robison, & Smith, 1972), and seriousness of prior offences (Steadman and Cocozza, 1974; Thornberry & Jacoby, 1979).

Although the utility of static variables has been recognized, the utility of dynamic variables has received much controversy. However, in a meta-analysis of 131 studies, Gendreau et al., (1996) examined predictor variables for adult recidivists. It was found that dynamic predictors performed as well as static predictors. The best predictor domains were: criminogenic needs (dynamic), criminal

history/history of antisocial behaviour (static), social achievement (dynamic), age/race/gender (static), and family factors (static).

It has been argued that the use of static variables does not permit reclassification (Gendreau et al., 1996). This is undoubtedly true, as static variables do not change. However, financial stressors on the penal system have brought about the ideation of the 'new penology' (Feeley & Simon, 1992). The 'new penology' moves the focus away from individual management and redirects it to the management of large aggregates of offenders. The emergence of the new penology has amplified the need for static predictors, in order to efficiently process large offender aggregates.

Methodological Issues

Self-Reports and the Use of Official Records

Many studies use only official records (e.g., Brennan, Mednick & John, 1989; Martinez, 1997). The use of official records however, does not reflect the true occurrence of violence (Monahan, 1981; Farrington, 1982) as much violence goes unreported. In an attempt to overcome this problem, researchers have used self-reports. However, self-reports are not without their problems. Jones (1993), after Farrington's 1985 study, questioned the validity of self-reports. Farrington found that the use of official records could produce a higher rate of accuracy than self-report measures. Self-reports are not only subject to under-reporting (Petersilia, 1978; Klassen & O'Connor, 1988b), but also over-reporting (Jones, 1993; Klassen & O'Connor, 1988b). The measurement of self-reports often involves a degree of subjectivity (Gendreau et al., 1996). These problems can be exacerbated by the fact

that self-reports often require accurate recall of offences over an extended period of time (Brennan et al., 1989).

Monahan (1981) and other researchers (Farrington, 1982; Jones, 1993) recommended the use of official records in conjunction with self-report methods to overcome such problems. Whilst this approach may alert the researcher to under-reporting, there is no real way of knowing if participants over-report. If participants are only required to report arrests or convictions, over-reporting may be checked against criminal records. However, such circumstances would negate the use of self-reports entirely.

Although the use of self-reports and official records both have disadvantages, to date there is no better alternative for the collection of data on violent behaviours. (This excludes domestic violence, as 'significant others' may be encouraged to report rates of violence.)

Whether the researcher chooses to use these methods solely or in conjunction with others depends largely on the chosen variables. For example, data for static variables would be more reliable if obtained from official records. Whereas, information for dynamic variables from official sources (even if available) would be less reliable, simply because of the mutable nature of dynamic variables.

Arrest, Conviction and Incarceration

Almost all prediction studies focusing on offending use some form of official records to measure levels of criminal activity. Offending levels are based on either arrest, conviction or incarceration. However, Jones (1993) stated that it was impossible to determine the extent to which such measures are confounded by the

criminal justice system. For example, selectivity or biases (e.g., by race or social status) of arrest procedures, may identify those factors as predictors (Jones, 1993).

Jones (1993) also stated that such problems were compounded when using convictions or incarceration as the level of measurement. Studies by Eggleston (1976) and Broadhurst (1987) suggested that selectivity and biases throughout the criminal justice system did affect arrest, conviction and incarceration. However, Lincoln and Wilson (1994) stated that the rate of Aboriginal imprisonment was 17 times higher than imprisonment for non-Aboriginals. Aboriginals were arrested at 29 times the rate of non-Aboriginals. These figures suggest that using arrest as the criterion, selectivity and bias have a greater effect on the outcome than if incarceration or conviction were the criterion measurement.

It is also true, as suggested by Jones (1993), that the workings of the criminal justice system (e.g., plea bargaining) may create a false impression of offence severity, by either upgrading or downgrading charges. However, Jones (1993) did not acknowledge that arrest does not confirm guilt. Furthermore, the use of the initial charge may create a false impression of dangerousness. For example, if an offender is initially charged with grievous bodily harm and the victim dies after the charge has been laid, the charge is upgraded to murder. Circumstances such as these have the effect of confounding offence severity.

Conviction therefore is as equally valid as arrest in the measurement of violent behaviour. However, it may be preferable that an individual's violent behaviour be measured by guilt (and level of offence severity) than by a premature accusation.

Statistical Methods

Receiver Operating Characteristic (ROC) curves were invented by Birdsall (Swets, 1986a) for use in signal detection theory. Recently Mossman (1994a, 1994b) suggested that the use of ROC curves for violence prediction was advantageous for two reasons. First they were unaffected by base rates as were previous actuarial methods. Secondly they were unaffected by biases in prediction outcomes as were clinical predictions (Mossman, 1994a, 1994b). For these reasons ROC curves have become the chosen statistic in violence prediction studies.

ROC curves assume the existence of the criterion, in this instance violent behaviour. The prevalence of the criterion within a specified population over a certain period of time is referred to as the 'base rate' (Borum, 1996). The base rate in any given population (e.g. the community, psychiatric ward or prison) does not remain static. The variation in base rate is due to covariates both within and between populations. Base rate variation makes it difficult to compare violence prediction studies (Mossman, 1994a). However, by using ROC curves, areas under the curves (i.e., the accuracy rate) of different models may be compared (Hanley & McNiel, 1982; DeLong, DeLong, & Clarke-Pearson, 1988).

The ROC curve is a visual representation of the trade-off of one criteria against another. For example, the ROC curve shows the probability of making choice A when it occurs, plotted against the probability of choosing A when B occurs (Swets, 1986b).

The power of the ROC curve is that two criteria can be mathematically determined. The trade-off between sensitivity and specificity determines the curve of the ROC. Sensitivity refers to the accuracy rate of true positives, whilst specificity is

the accuracy rate of true negatives (Swets, 1986a). In the present context, sensitivity refers to the likelihood of a violent offender who commits further violent offences being correctly identified. Whilst specificity refers to the likelihood of a violent offender who does not commit further violent offences being correctly identified.

The area under the curve (AUC) is the probability that a randomly chosen offender that committed a further violent offence will have a higher score than a randomly chosen offender who did not commit a further violent offence (Rice & Harris, 1995a). The AUC may be calculated using a variety of methods. For example, Wilcoxon, the Trapezoidal method, (Hanley & McNiel, 1982) or the Common Language effect (Rice & Harris, 1995a) are three ways for calculating the AUC.

By maximizing sensitivity and specificity, the rate of misclassification can be minimized. Thus the overall cutoff point for the scale can be manipulated to achieve the maximum amount of both true positives and true negatives. Although the error rate can be minimized, it must be noted that the cutoff point is finally determined by policy.

The Present Study

With the large aggregate of violent offenders being received at prisons, it is impossible to target them all for behavioural intervention programs. As a result it has become necessary to develop a method for selecting those most in need of intervention. The developers of the Violent Offender Treatment Program (VOTP) (Howells, Watt, Hall, & Baldwin, 1997) have devised an instrument designed to screen high-risk violent offenders for intervention therapy.

The instrument is intended only as a screening process. As a result of the screening process, some offenders will be directed away from the program, and others will be directed towards the program. For those directed towards the program, full assessment will then be administered to ensure program suitability. However, as some offenders will be directed away from the program, it is important to identify those most at risk of committing future violent offences. Hence the development of the VOTP RAS.

The following study was undertaken to (a) determine the reliability of the VOTP RAS, (b) assess the predictive accuracy of the VOTP RAS, (c) determine the optimal cutoff point, (d) assess instrument accuracy over varying time-at-risk periods,(e) examine cutoff points for the varying time-at-risk periods, and (f) examine variable utility.

Method

Design

The present study was a longitudinal retrospective case analysis. The predictor variable was the score obtained from the VOTP RAS. The criterion or outcome variable was future violent offences.

Participants

Participants for this study comprised violent male offenders aged 18 years and over (M = 28 yr. at time of release). A quota sample of 202 offenders released from January 1985, who had served an unbroken minimum prison sentence of 1 year for

offences which included at least one violent offence were chosen. The sample consisted of 50 aboriginals and 152 non-aboriginals. Participants were released from the index offence between January 1985 and up to and including July of 1987.

As criminal histories were being examined, confidentiality was ensured by coding participants. The master sheet which gave the participant's name and matching code remained in the possession of the Alternatives to Violence Unit at the Ministry of Justice of Western Australia.

Raters

The VOTP RAS was rated by two prison officers from the Sentence Planning section at Casuarina Prison, and the researcher. Neither of the prison officers had previously used the RAS. Ratings on the adapted version of the VOTP RAS were also completed by two Casuarina Prison Officers from Sentence Planning and the researcher. None of the raters received training in the completion of the VOTP RAS.

Materials

Preliminary ratings used the original version of the VOTP RAS with original instructions (Appendix A). The VOTP RAS comprised seven items, five of which related to previous and current criminal offences. The remaining two items related to alcohol and drug misuse. Items on the VOTP RAS were unevenly weighted. For preliminary ratings, the RAS was applied to ten current sentence plans. Final ratings were completed using the adapted RAS (Appendix B). The adapted version was applied to 20 court histories of offenders released in 1987.

For the study, the adapted VOTP RAS and the Subsequent Offence Form (Appendix C) were applied to court histories. The Subsequent Offence Form consisted of three items. The first item related to subsequent offence type (e.g., violent or nonviolent), and the other two items related to first violent and nonviolent failure.

Sections from the Western Australian Criminal Code (1983) dealing with violent offences, were grouped according to level of severity (Appendix D). Groupings were used as a basis for scoring the level of convicted offence.

Procedure

Violent convictions prior to, and including, the index offence were scored on the VOTP RAS according to groupings of the Criminal Code. Traffic offences were used only to obtain information for alcohol and drug items. Possession offences were assumed to indicate the use of the drug in question. Drugs were assumed to be taken orally unless the offence related to possession of injecting equipment (e.g., needles). When a number of convictions arose from a single offence, only the most severe conviction was recorded.

Each offenders' convictions for a 10 year period following the index offence (e.g., released June 1985, followed up to and including June 1995) were recorded on the Subsequent Offence Form.

Results

The results will be presented in three sections. The first section shows the results of the inter-rater reliability analysis. The second section shows the Receiver Operating Characteristic curves and the corresponding cutoff points. The third section shows the correlations between variables and the total score, and the inter-item correlations.

Inter-rater reliability

Rater (R) agreement for 10 criminal records were analysed using kappa (κ). The κ values for the preliminary rater agreement were R1 * R2 $\kappa = 0.64$, R1 * R3 $\kappa = 0.74$, and R2 * R3 $\kappa = 0.54$. Overall mean agreement was $\kappa = 0.64$. Kappa values for final rater agreement of 20 court histories are shown on Table 1. An overall mean agreement of $\kappa = 0.82$. was attained.

Table 1.

Kappa values for rater agreement.

Raters	Kappa
R1 * R2	.85
R1 * R3	.81
R2 * R3	.80

Receiver Operating Characteristic curve

A ROC curve was computed for the 10 year risk period, using the statistical program ROCFIT (Metz, Shen, Wang, & Kronman, 1989). To meet ROCFIT

requirements, the 20 observed categories were collapsed into 9 working categories (Appendix E), with the required minimum of 5 cases per category (Hanley & McNiel, 1982). Weighted rank order data assumed interval level data when totaled. Calculated operating points (Appendix F) were used as Expected Operating Points (Appendix G) were subject to variation due to groupings. The area under the curve (AUC) was 0.76 (*SD* = 0.03). Figure 1 shows the constructed ROC curve for the 10 year time at risk period.

Cutoff point

The sample base rate (BR) was 47%. The 2x2 contingency tables (Appendix H) show that the cutoff point which maximised correct classification and minimised error was 11 (i.e.,< 12). The contingency table for the cutoff point of 11 is presented on Table 2. At the cutoff point of eleven, 141 (70%) cases were correctly classified, resulting in a sensitivity of 70% and a specificity of 69% (False positive fraction = .31).

Table 2.

Classification for cutoff point 11 (< 12) at the 10-year at-risk period.

TP	FP	FN	TN
33.17% (67)	13.86% (28)	16.33% (33)	36.63% (74)

Accuracy for varying time at risk intervals

To examine predictive accuracy of the VOTP RAS over varying time-at-risk periods, ROC curves for 36 month, 60 month and 84 month intervals were computed.

For the 36 month time at risk, 20 observed categories were collapsed into 7 working categories. The success group consisted of 140 cases, 62 cases had failed (BR = 31%).

The AUC for the 36 month time at risk period was 0.73 (*SD* = 0.04). Figure 2 shows the constructed ROC curve for the 36 month time at risk period. A corresponding cutoff point for the 36 month time at risk period was 15. This resulted in 148 correctly classified cases (specificity = .82, sensitivity = .51). However, if a cutoff point of 11 was retained, a higher sensitivity of .74 was achieved, resulting in a specificity of .61. A total of 132 cases were correctly classified for the cutoff point of 11. Table 3 shows the classification of participants for the cutoff points of 15 and 11 at the 36 month interval.

Table 3.

Classification at the cutoff points of 15 and 11 at the 36 month interval.

TP	FP	FN	TN	
15.84% (32)	14.85% (30)	11.88% (24)	57.43% (116)	Cutoff 15 at 36 month interval
TP	FP	FN	TN	
22.77% (46)	7.92% (16)	26.73% (54)	42.57% (86)	Cutoff 11 at 36 month interval

At the 60 month time at risk interval, 74 cases had failed and 128 cases had succeeded (BR = 37%). The 20 observed categories were collapsed into 8 working

categories. A ROC curve was computed, resulting in 7 expected operating points. Figure 3 shows that the AUC at the 60 month time at risk interval was 0.74 (*SD* = 0.03). A corresponding cutoff of 15 resulted in 142 cases correctly classified. A cutoff of 15 showed sensitivity at .47 and specificity at .83. For a cutoff of 11, sensitivity was .80 and specificity was .64. At the cutoff point of 11, 138 cases were correctly classified. Table 4 shows the classification for the cutoff points of 15 and 11 at the 60 month time-at-risk interval.

Table 4.

Classification for cutoff points 15 and 11 at the 60 month interval.

TP	FP	FN	TN	
17.33% (35)	19.31% (39)	10.4% (21)	52.97% (107)	Cutoff 15 at 60 month interval

TP	FP	FN	TN	
27.23% (55)	9.4% (19)	22.28% (45)	41.1% (83)	Cutoff 11 at 60 month interval

At the 84 month time at risk interval, there were 116 cases of success and 86 cases that had failed (BR = 43%). The 20 observed categories were collapsed into 10 working categories. The plotted ROC curve is shown in Figure 4. The AUC for the 84 month interval was 0.72 (*SD* = 0.03). A corresponding cutoff point of 11 minimised error (60 cases incorrectly classified), resulting in a sensitivity of .73 and a specificity of .68. Table 5 shows the cutoff point of 11 for the 84 month time-at-risk interval.

Table 5.

Classification for cutoff of 11 for the 84 month interval.

TP	FP	FN	TN	
31.19% (63)	11.39% (23)	18.32% (37)	39.11% (79)	Cutoff 11 at 84 month interval

Variable Utility

Using the Statistical Package for Social Sciences (SPSS), Spearman correlation coefficients (r) were computed between each variable and the RAS score. Table 6 shows values and significance level for each variable.

Table 6.

Spearman correlation coefficients for each variable.

Variable	r	Sig.
Current offence	0.34	.000**
Offence severity	0.81	.000**
PVO	0.87	.000**
PNVO	0.63	.000**
Age at first offence	0.37	.000**
Alcohol offences	0.62	.000**
Drug offences	- 0.09	.200

$N = 202$. $p < .001$**. PVO = Past Violent Offences. PNVO = Past Nonviolent Offences.

All values were above .3 except item 7, drug offences, which showed a negative correlation of - 0.09 which was not significant ($p > .05$). Spearman Correlation Coefficients for item by item are shown on Table 7.

Table 7.

Item by item coefficients

	Current	Severity	PVO	PNVO	Age	Alcohol
Severity	.23**					
PVO	.23**	.90**				
PNVO	.01	.36**	.41**			
Age	-.02	.18*	.18*	.41**		
Alcohol	.18*	.4**	.48**	.31**	.02	
Drug	.34**	-.23**	-.23**	-.03	.01	-.25

Current = current offence. Severity = most serious offence. PVO = previous violent offences. PNVO = previous nonviolent offences. Age = age at first offence. Alcohol = alcohol related offences. Drug = other drug misuse. $p < .001^{**}$. $p < .05^{*}$.

The extremely high correlation between Most serious offence and Past violent offences indicated the redundancy of one of those variables. All variables were significantly correlated with at least one other variable.

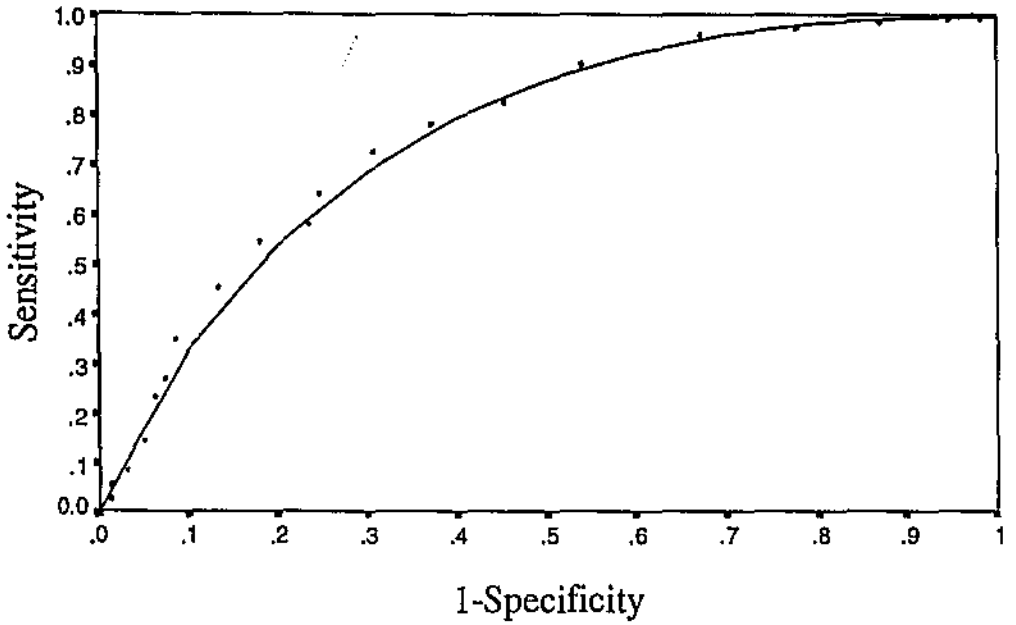

Figure 1. The ROC curve showing cutoff points for 10-year time-at-risk period.

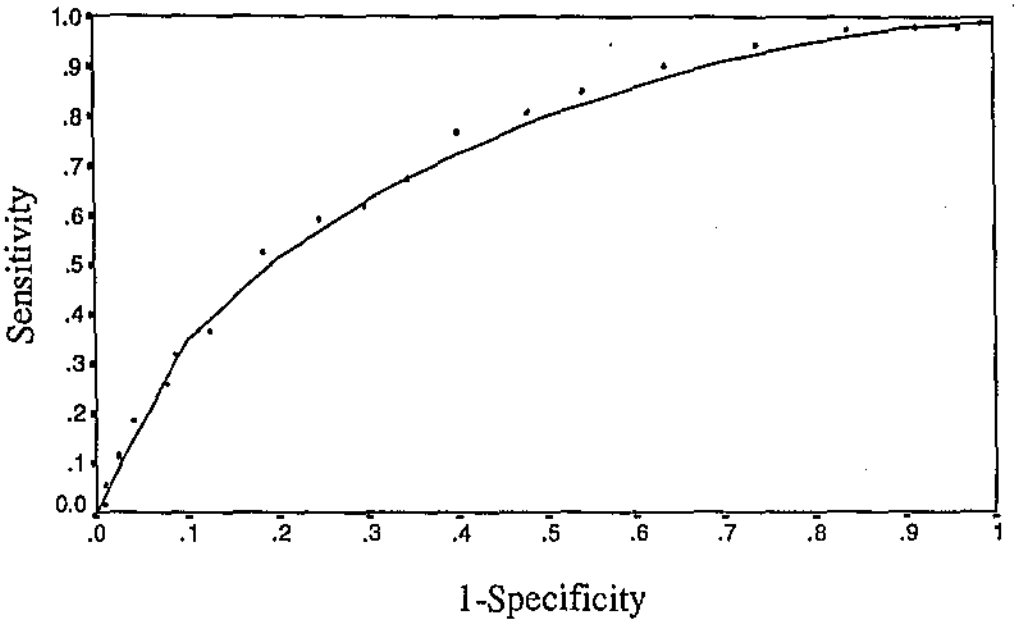

Figure 2. ROC curve showing cutoff points for 36 month time-at-risk period.

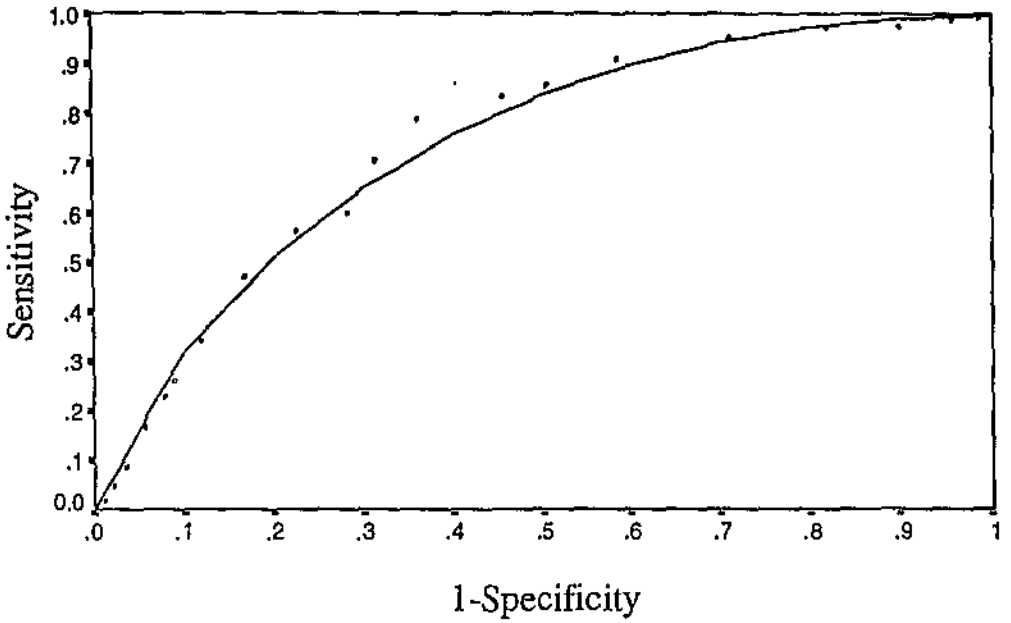

Figure 3. ROC curve showing the cutoff points for 60 month time-at-risk period.

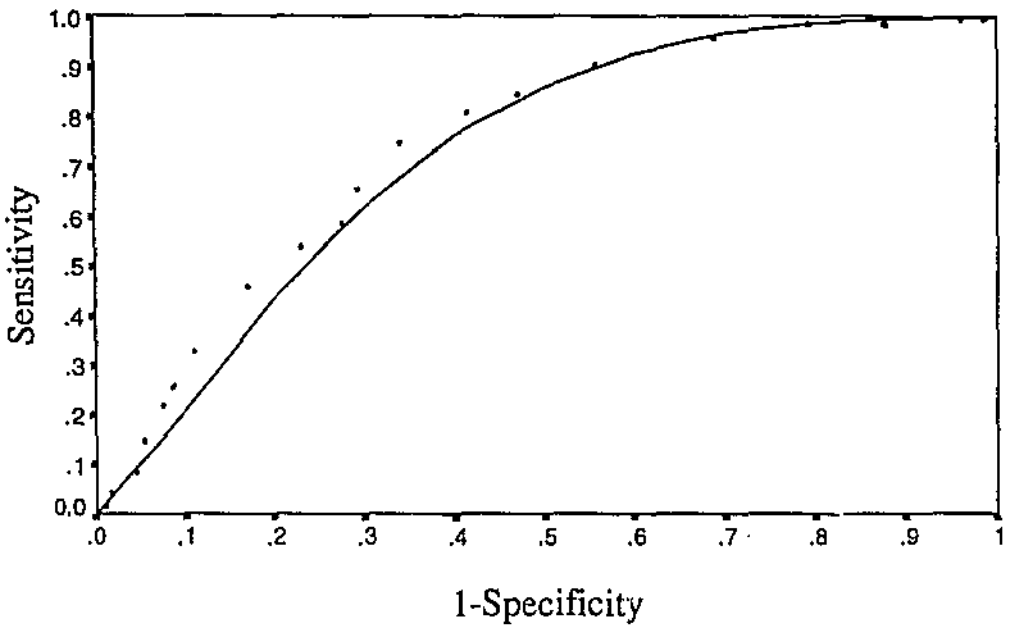

Figure 4. ROC curve showing the cutoff points for 84 month interval.

Discussion

The present study examined (a) the reliability of the VOTP RAS, (b) the predictive accuracy of the VOTP RAS, (c) the optimal cutoff point, (d) accuracy over varying time-at-risk periods,(e) cutoff points for the varying time-at-risk periods, and (f) variable utility. The accuracy rate of the VOTP RAS indicated that violence could be predicted accurately enough to aid decision-making with respect to program entry. The accuracy of classification of violent and nonviolent offenders by the VOTP RAS was congruent with the study by Rice and Harris (1995).

Predictive Accuracy

The VOTP RAS ROC curve for the 10-year time-at-risk period showed a higher rate of accuracy than the SIR scale (Bonta, Harman, Hann & Cormier, 1996) which had an accuracy rate of 72% for general recidivism. However, the accuracy rate of the SIR scale for violent recidivism was considerably lower. Narrowly-defined violence (65%) elicited an accuracy rate marginally higher than the broadly-defined violence (64%).

The VOTP RAS also rated favourably in comparison to McNiel and Binder's (1994) instrument. Developed for short-term predictions, the Screening Checklist (McNiel & Binder, 1994) had an overall accuracy rate of 65.4%. The VOTP RAS showed an improvement over the Screening Checklist of approximately 10%. Furthermore, the Screening Checklist had a high false positive rate (42.3%) in comparison to the lower false positive rate of the VOTP RAS (13.86%).

The instrument used by Rice and Harris (1995) had a correct classification rate of 70% and a predictive accuracy of 76%. The Rice and Harris (1995) instrument yielded a false positive rate of 21%. The VOTP RAS also showed a correct classification of 70% and a predictive accuracy of 76%. The false positive rate for the VOTP RAS was 13.86%, approximately 8% lower than that of Rice and Harris' (1995) instrument.

Classification rates for the VOTP RAS were lower than those found by Klassen and O'Connor (1988a, 1988b). However, the correct classification of violent offenders in Klassen and O'Connors (1988a) study was only 59.3%. In Klassen and O'Connor's (1988a) first study, the correct classification rate was 85.3% with a high false positive rate of 40%. In their second study (Klassen and O'Connor, 1988b) the correct classification rose to 88.1%, whilst the false positive rate dropped to 17.6%. However, as Klassen and O'Connor (1988a) used 'tailor-made' (Klassen & O'Connor, 1988b) statistical techniques, the high classification may be an artifact of the techniques used. Furthermore, the use of 'tailor-made' statistical techniques indicates that the study cannot be replicated.

In comparison to base rates, the VOTP RAS again yielded a higher predictive accuracy. For example, the sample base rate for the ten-year time-at-risk period was 47% whilst the accuracy of the VOTP RAS was 76%. This shows an improvement of 29%. For 34 month, 60 month and 84 month time-at-risk intervals the base rate was 31%, 37% and 43% respectively. The predictive accuracy for the time-at-risk intervals was 73%, 74% and 72% respectively.

These findings were congruent with those of Rice and Harris (1995). They found that for varying time-at-risk periods (3.5, 6, and 10 years) their instrument's

accuracy remained between .74 and .75. Although accuracy rates appear modest, they are some of the highest achieved in the area of violence prediction.

As ROC curves are not sensitive to base rates, it was expected that the accuracy rate would remain relatively stable over time. It was not expected that the lowest accuracy rate should occur at the 84 month time-at-risk interval. However, the lower accuracy rate may be a due to a lull in the commission of violent offences, whilst nonviolent offences continued to be committed. This would indicate that sequential violence specialization (Brennan et al., 1989) may be less prevalent than distributional violence specialization. The task of correctly identifying distributional 'specialists' would be more difficult than identifying sequential 'specialists'. For example, the previous offence of a sequential specialist would be indicative of the next offence being violent. Whereas, distributional specialists would have to commit a number of offences before they could be identified as violence specialists.

Many studies on violent offending are conducted within a psychiatric population. Due to concern for liberty issues, such studies emphasize the ramifications of the false positive rate (e.g., Klassen & O'Connor, 1988a, 1988b; Rice & Harris, 1992, 1995a, 1995b; McNiel & Binder, 1994). However, the present study is more concerned with the false negative rate. Although false positives are of concern, they will be either identified as true positives through full assessment, or screened out. However, this is not the case with false negatives. As false negatives will be screened out using the VOTP RAS, they will not reach the point of full assessment.

The misclassification of false negatives means that a percentage of violent offenders in need of behavioural intervention will not get the opportunity to

participate in an intervention program. However, in the event of a person identified as a false negative violently reoffending, it is likely that on their next prison admission their RAS score would be higher. This would then give them the opportunity to participate in a violence intervention program.

Public Policy and Cutoff Points

The cutoff points that minimized misclassification for the 36 and 60 month intervals were equal to or less than 15. However, a cutoff of 15 resulted in extremely poor sensitivity. The lack of sensitivity indicated that many violent offenders would not be correctly identified. Misclassification of violent offenders' would mean that those most in need of behavioural intervention would be directed away from intervention programs. At the cutoff point of 11 however, the sensitivity was raised to an acceptable level. For these reasons it is recommended that the cutoff point remain at 11 for all time-at-risk periods. Although the cutoff point of 11 improved the sensitivity of the test, it also increased the rate of false negatives.

However, the cutoff point may be altered according to public policies. If violent behaviour became a more salient policy issue, it is expected that the cutoff point would be lowered. This would have the effect of correctly identifying more violent offenders. By lowering the cutoff point, it would increase the aggregate of offenders to be fully assessed. However, although many false positives would be screened out through full assessment, and many false negatives would be identified as true positives, the costs involved with lowering the cutoff point may not be financially beneficial.

Instrument Reliability

The preliminary inter-rater agreement was considered low at $\kappa = .63$. In an effort to enhance rater reliability a number of items were adapted. As prison officers were to administer the VOTP RAS, item 1(c) and 2(d) were changed from 'injuries life threatening' to 'grievous bodily harm'. Item 6 (use of alcohol) was changed from type of drinker (e.g., occasional user, binge drinker, moderate heavy user, regular heavy user) to number of alcohol-related offences (e.g., 1 alcohol offence, 2 alcohol offences, etc.).

Two subitems of item 7 (other drug misuse) were also altered. Item 7(b) was changed from 'occasional user, non-intravenous' to 'cannabis'. Item 7(c) was changed from 'moderate-heavy user, non-intravenous' to 'other non-intravenous drugs'. Changes on items 1(c) and 2(d) were intended to make the instrument easier for staff in the legal system (i.e., prison officers) to score. Changes for items 6 and 7 were necessary for two reasons. First there was a lack of information in criminal files. Second the level of subjectivity required to score items 6 and 7 reduced reliability.

The adapted VOTP RAS attained an inter-rater reliability similar to that achieved by Harris, Rice and Quinsey (1993). It remains unclear whether the authors (i.e., Harris, Rice and Quinsey) completed the ratings or whether other people completed the ratings. However, as diagnostic criteria were used as variables, it is inferred that ratings were completed by Harris et al., (1993). If completed by the authors, it is equivocal as to whether the instrument would achieve the same level of rater reliability when used by others.

Values reaching .80 appear to be higher than most (e.g., McNiel & Binder (1994) had a rater agreement of $\kappa = .75$). However, Klassen and O'Connor's (1988a) instrument attained a rater agreement over 95%. Once again as the authors were the only two raters, it is considered that the instrument may not achieve the same level of reliability when used by others. Studies in which instruments are rated by those who developed them may not be replicable. This may reduce the reliability and hence the validity of the instrument.

Variable Utility

Past violence was found to have the highest correlation with the total score. Offence severity also had a strong correlation with future violence. However, the item-by-item analysis showed that the past violence and offence severity items were strongly correlated. This suggests that they may have been measuring the same construct. The high correlation between offence severity and past violence suggests redundancy of either variable. As past violence had a higher correlation with the total score, it is considered that the offence severity item is redundant.

Changes to the drug and alcohol items were expected to minimally affect their utility. While the changes seem to have negated the effect of the drug item, the alcohol item retained a statistically significant moderate correlation. The use of alcohol before or during the commission of a violent offence has consistently been found to be predictive of future violence (e.g., Klassen & O'Connor, 1988a; Harris & Rice, 1993). However, static predictors of alcohol use have also been shown to have predictive utility in assessing future violence (Webster et al., 1994).

Static Vs Dynamic Variables

It has been argued by some (e.g., Gendreau et al., 1996) that static variables do not allow for reclassification or rehabilitation. However, as the VOTP RAS is intended as a screening instrument for violent offenders, it is unlikely that the VOTP RAS will be applied to offenders sentenced for crimes other than violence. Use of the RAS for offenders received for nonviolent crimes may be biased. For example, the first item on the VOTP RAS relates to the level of violence in the current crime. As the recommended cutoff point is 11 and item one can generate up to 5 points towards the cutoff, it is considered that most offenders whose current offence is not violent will not attain a RAS score of 12 or more points.

Actuarial Vs Clinical Predictions

Actuarial methods evaluating offender violence have consistently been shown to be more accurate than clinical predictions (Gardner et al., 1996a; Steadman & Cocozza, 1974; Thornberry & Jacoby, 1979). A landmark case in the accuracy of clinical predictions was the Baxstom case (Steadman & Cocozza, 1974). Baxstrom, a 'mentally ill' inmate served out his sentence in a hospital prison. Upon sentence completion, Baxstrom was civilly committed without judicial review. Although the practice of civilly committing 'mentally ill' inmates after sentence completion was not uncommon, Baxstrom took his case to the United States Supreme Court. After receiving papers from Baxstrom, his counsel, Louis B. Polsky wrote:

> After reading through the papers I realized, to my amazement, that Baxstrom, supposedly insane, had managed to meet all of the highly

> technical procedural steps for timely getting his case from the County Court, to the Appellate Division, to the Court of Appeals, to the U. S. Supreme Court.
>
> Steadman & Cocozza (1974, p.46).

Not surprisingly, Baxstrom won his case and as a result, Johnnie Baxstrom and 966 other patients were transferred from the prison hospital to 18 civil hospitals. Although over half of the Baxstrom patients were later released (including Baxstrom himself), Steadman and Cocozza's sample consisted of 98 released Baxstrom patients. Although the Baxstrom patients were considered some of the most dangerous in the country, Steadman and Cocozza found that only 20 Baxstrom patients were rearrested, only seven of the 20 had been reconvicted of violent offences.

Researchers Thornberry and Jacoby (1979) had the opportunity to replicate the Steadman and Cocozza (1974) study in the Dixon case. In *Dixon V. Attorney General of the Commonwealth of Pennsylvania*, the practice of civil commitment after sentence completion without benefit of judicial review was again deemed legally unconstitutional.

Of the 586 Dixon patients that were transferred to civil hospitals, 107 of which were considered "too dangerous" to be released to civil hospitals (Thornberry & Jacoby, 1979, p.22) 414 were later released. Of the released patients, only 98 were rearrested. Of those classified as dangerous, only 46 of 60 were rearrested for violent offences (ranging from threats to murder). However, only 14 were serious enough for rehospitalisation.

Results of the Baxstrom and Dixon cases were remarkably similar. Of the Baxstrom patients 14.3% were subsequently considered dangerous. Similarly, 14.5% of the Dixon patients were subsequently found to fit the dangerousness criteria. In the Baxstrom and Dixon cases 80 to 86 percent of predictions proved to be wrong (Monahan, 1981). This showed the over-prediction of dangerousness by mental health professionals.

Implications

Many risk instruments incorporate other scales (e.g., Wenk, Robison & Smith, 1972) some of which are diagnostic (e.g., Harris et al., 1995). The incorporation of other scales into risk instruments often makes them lengthy and time-consuming to administer. These problems are compounded when the instrument includes a diagnostic scale which requires a psychologist or psychiatrist to interpret results. Fiscal resources are simply not available to supply a clinician to select offenders for evaluation for the entry into violence treatment programs.

It is considered that the VOTP RAS is a valuable tool for the 'new penology' described by Feeley and Simon (1992). The use of the VOTP RAS simplifies the task of screening large offender aggregates for risk of future violence. The brevity of the VOTP RAS assists in its easy application. As no information other than that available from court histories is needed to score the RAS, it is an economically viable instrument to use in comparison to other risk instruments. As no diagnostic items are included in the VOTP RAS, prison officers are able to apply it to offender records. The advantages of the VOTP RAS are a reduction in the time, money and expertise needed to score other instruments. Resources saved in the assessment of risk, could

be redirected into other areas such as improving program delivery and program evaluation.

The use of actuarial instruments has increased the accuracy of violence predictions (Monahan & Steadman, 1994; Gardner et al., 1996a). Successfully identifying violent offenders using the VOTP RAS then allows those offenders to receive behavioural intervention to reduce violent behaviour. This is particularly salient in light of results from the study by Brennan et al., (1989) which showed sequential and distributional violence specialization. As a relatively small aggregate of violent offenders were responsible for a relatively large amount of violent crimes, the correct identification of violence 'specialists' could impact on the penal system in a number of ways.

For example, crimes of violence receive the longest sentences (e.g., murder, armed robbery). Lengthy sentences incurred for violent crimes increases prison overcrowding and drains fiscal resources. If violence specialists were correctly identified the cost both in human (victim injury) and economical terms could be reduced. The effective treatment of violent offenders would also increase public safety, and possibly the public perception of safety.

The VOTP RAS and the Risk Principle

The VOTP RAS has demonstrated a relatively accurate classification rate of low-risk and high-risk offenders (76%). The ability to discriminate between low- and high-risk offenders may have a great impact on treatment effectiveness. As demonstrated by Andrews and colleagues (Andrews, Kiessling et al., 1986: Andrews,

Robinson, & Balla, 1986; Andrews, Bonta & Hoge, 1990; Andrews, Zinger et al., 1990), treating low-risk offenders may have a minimal or detrimental effect.

In contrast, the delivery of treatment to high-risk cases is considered to have a beneficial impact on offenders (Andrews, Zinger et al. 1990). In applying the risk principle to violent offenders, the benefits would not only be gained by the offenders but also their victims, the penal system and the health system (i.e., for injuries sustained by either the offender or victim). However, identifying high-risk offenders is simply the first step in delivering effective treatment. As Andrews and colleagues (Andrews & Bonta, 1994; Andrews, Zinger et al., 1990) pointed out, other aspects such as needs and treatment matching are integral factors for the effective delivery of treatment.

Treatment matching refers to the matching of client learning-style, need and therapist (Andrews, Kiessling et al., 1986; Andrews, Zinger, et al., 1990). For example, some clients may respond better to a cognitive-behavioural approach while others may respond better to a psychotherapy approach. Similarly, some clients may be more open to a certain therapists' style. As Andrews, Zinger, et al., (1990) stated, the three principle considerations for effective treatment were risk, needs, and responsivity (i.e., treatment matching).

Treatment matching is imperative if programs are to be effective. If violent offenders do not receive intervention, it is unlikely that established behavioural patterns will change. Similarly, established behavioural patterns are unlikely to change if violent offenders receive ineffective treatment (Andrews, Zinger, et al., 1990). However, effective treatment reduces violent behaviour. This in turn reduces violence both in prison and in the community and reduces prison over-crowding.

The Future of Risk, Violence and ROC Curves

In their 1992 article, Mossman and Somoza suggested a futuristic scene (the year 2014) in which violent offenders could be discriminated from nonviolent offenders by a biological test. However, as the accuracy of tests are rarely perfect, Mossman and Somoza (1992) suggested the use of ROC curves for the analysis of data.

Researchers such as Virkkunen et al., (1989) have begun the search for psychobiological variables that play a role in violence. Whilst initial results appeared promising (84% correct classification on the basis of blood glucose nadirs and 5-hydroxyindoleacetic acid), further studies have not been able to replicate results (DeJong et al., 1992). However, if such results are later replicated, ROC curves offer a way of integrating biological and psychological predictors. Furthermore, optimal levels of the integration of both biological and psychological predictors could be determined using ROC curves. As a result, biological predictors may partially validate psychological predictors.

DeJong et al., (1992) offered some explanations as to why results from the psychobiological study (Virkkunen et al., 1989) did not concur with earlier studies. However, statistical methods used to analyse data were not considered. The statistical methods used in violence prediction studies have been notoriously unreliable. As most statistical techniques are sensitive to base rates, the accuracy of a model varies even within the same population (Mossman, 1994a). Many researchers (e.g., Martinez, 1997; Klassen & O'Connor, 1988a, 1988b; Dejong et al., 1992) have used discriminant analysis or similar statistics. Whilst discriminant analysis may be a

useful statistic for determining the contribution of each individual variable, it tells little about the accuracy rate of the overall model.

Furthermore, discriminant analysis offers no effective way of choosing one criterion over another when the criteria overlap. However, the overlapping of criteria is a reality of violence prediction. As ROC curves were (a) designed specifically for the purpose of optimizing decision-making of overlapping criteria, and (b) not sensitive to base rates, they are the most accurate method for testing models.

Limitations

As the instrument was scored using groupings derived from the *Criminal Code of Western Australia* (1983) it may not be readily transferable to other geographical locations. However, as most other jurisdictions have the offence grievous bodily harm or an equivalent offence, it is considered that with caution other jurisdictions could also use the VOTP RAS as a tool to aid entry to treatment programs for violent male offenders.

However, before the VOTP RAS was utilized in other areas, there are a number of considerations to be addressed. First results may be an artifact of the cohort group (Martin, 1996). Cohort effects may be the result of intervention strategies or policy decisions current in Western Australia preceding offenders' release. Secondly, other Australian states may have a broader multicultural offender base than Western Australia.

As the studied sample may not be representative of other Australian prison populations, and only a single cohort was used, results should be interpreted cautiously. The correct classification rate may be sample-specific. Before the VOTP

RAS was utilized in similar populations universally, it is recommended that predictors be examined to determine suitability to the intended population, as factors such as cultural influences may affect the efficacy of predictors. Furthermore, item weightings may be affected by either cultural or geographical locations.

Attrition Rate

One of the problems that researchers often face when using official records (e.g., arrest, conviction, or incarceration records) is the unknown drop-out rate. Offenders may die, move interstate or leave the country. Whilst this is not considered to greatly affect the sample, it may play some part. The unknown attrition rate may elevate the number of either false positives or true negatives. For example, if an offender with a history of violent offences left the state, there would be no record of further offences. Since there was no record of further offences, if the offender had received a high RAS score, they would be considered a false positive. Similarly, if an offender had received a low RAS score, and had no record of further offences, they would be categorized as a true negative. However, it would be unknown whether the lack of further recorded offences was due to no further offences being committed, or convictions being recorded elsewhere.

Some researchers (e.g., Martinez, 1997) have overcome this issue, to some extent, by using centralized files (e.g., Federal Bureau of Investigation files). Whilst this may reduce the unknown attrition rate, it does not necessarily resolve the issue. For example, FBI files may not offer information as to whether the offender has left the country. However, files in Australia have not been federally centralized.

The attrition rate may have had some effect on the final accuracy rate (Coolican, 1994), as there were several offenders who had extremely high scores, whose long criminal careers ended abruptly. However, such cases could not be excluded simply because their records contained no further offences. In the event that these offenders had died or left the state, the rate of false positives would be artificially elevated. However, this would result in a conservative estimation of the instruments accuracy. This effect may also have influenced the true negative rate, as offenders who received low scores may have offended in another state or country. However, as the instrument's accuracy remained relatively high, the attrition rate is considered to only have had a minimal effect.

Rehabilitation

In addition to attrition, rehabilitation may have affected the accuracy rate (Coolican, 1994). The base rate may have been higher if offenders had not undergone some form of behavioural intervention. For example, during the past 10 years the Western Australian Ministry of Justice has been delivering the Skills Training and Aggression Control program to violent offenders. The program process may have, in some cases, been successful. This would have the effect of lowering the rate of intervening violent offences. However, as variables on the VOTP RAS are static, offenders may have attained high VOTP RAS scores but may not have violently reoffended during the follow-up period.

Further Research

Although the VOTP RAS is considered to have enough accuracy to be used in its current form as an aid to decision-making in relation program entry, a number of changes may improve its rate of accuracy. The following presents a number of suggestions which may enhance the accuracy of the VOTP RAS.

The present research was conducted on a representative sample of West Australian violent male offenders released between 1985 and 1987. Due to the relatively large proportion of aboriginal people in the sample, it is considered that the accuracy of the VOTP RAS should be examined in respect to specific populations. While it is considered that the overall accuracy would not change dramatically, the cutoff point for an aboriginal population may be higher or lower than the cutoff point for nonaboriginals.

Other possible changes which may enhance accuracy of the VOTP RAS is the exclusion of items 7 (drug misuse) and 2 (offence severity). These two items appear to add nothing, and item 7 may possibly detract from the accuracy of the VOTP RAS. However, the inclusion of an item pertaining to current age may be beneficial, as this variable has consistently been shown to be associated with risk (e.g., Steadman & Cocozza, 1974; Thornberry & Jacoby, 1979; Webster et al., 1994).

As this study was intended purely as a calibration study of the VOTP RAS, it is recommended that a prospective study using a larger sample of violent offenders be undertaken. It is envisaged that offenders be assigned as either high- or low-risk in accordance with the recommended cutoff point.

As concurrent validity was not examined in this study, it is recommended that the VOTP RAS be cross-validated with an already proven instrument. However, as

no other currently existing instrument specifically addresses violent offending in prison populations, an equivalent may have to be used. For example, the SIR was designed specifically for general recidivism. However, as the Bonta et al., (1996) study showed, the SIR had some success with both broadly-defined and narrowly-defined violent recidivism.

References

Andrews, D. A., & Bonta, J. (1994). *The Psychology of Criminal Conduct.* Cincinnati: Anderson Publishing Company.

Andrews, D. A., Bonta, J., & Hoge, R. D. (1990). Classification for effective rehabilitation: Rediscovering psychology. *Criminal Justice and Behavior, 14,* 19-52.

Andrews, D. A., Kiessling, J. J., Robinson, D., & Mickus, S. (1986). The risk principle of case classification: An outcome evaluation with young adult probationers. *Canadian Journal of Criminology, 28,* 377-386.

Andrews, D. A., Robinson, D., & Balla, M. (1986). The risk principle of case classification and the prevention of residential placements: An outcome evaluation of share the parenting. *Journal of Consulting and Clincial Psychology, 54,* 203-207.

Andrews, D. A., Zinger, I., Hoge, R. D., Bonta, J., Gendreau, P., & Cullen., F. T. (1990). Does correctional treatment really work? A clinically relevent and psychologically informed meta-analysis. *Criminology, 28*(3), 369-404.

Baird, S. C. (1981). Probation and parole classification: The Wisconsin model. *Corrections Today, May-June,* 36/38-41.

Blackburn, R. (1993). *The Psychology of Criminal Conduct.* Chichester: Wiley.

Bonta, J., Harman, W. G., Hann, R. G., & Cormier, R. B. (1996). The prediction of recidivism among federally sentenced offenders: A re-validation of the SIR scale. *Canadian Journal of Criminology, 38*(1), 61-79

Bonta, J., & Motiuk, L. L. (1985). Utilization of an interview-based classification instrument: A study of correctional halfway houses. *Criminal Justice and Behaviour, 12*(3), 333-352.

Bonta, J., & Motiuk, L. L. (1987). The diversion of incarcerated offenders to correctional halfway houses. *Journal of Research in Crime and Delinquency, 24*(4), 302-323.

Bonta, J., & Motiuk, L. L. (1990). Classification to halfway houses: A quasi-experimental evaluation. *Criminology, 28*(3), 497-506.

Bonta, J., & Motiuk, L. L. (1992). Inmate classification. *Journal of Criminal Justice, 20*, 343-353.

Borum, R. (1996). Improving the clinical practice of violence risk assessment: Technology, guidelines, and training. *American Psychologist, 51*(9), 945-956.

Brennan, P., Mednick, S., & John, R. (1989). Specialisation in violence: Evidence of a criminal subgroup. *Criminology, 27*(3), 437-453.

Broadhurst, R. G. (1987). Imprisonment of the Aborigine in Western Australia, 1957-1985. In K. M. Hazlehurst (Ed.), *Ivory Scales: Black Australia and the Law*. Sydney: New South Wales Press.

Broadhurst, R. (1991). Evaluating imprisonment and parole: Survival rates or failure rates? In S. McKillop (Ed.), *Keeping People out of Prison* (pp. 153-181). Canberra: Institute of Criminology.

Broadhurst, R. & Maller, R. A. (1990). Recidivism of prisoners released for the first time. *Australian and New Zealand Journal of Criminology, 23*(2), 88-104.

Brown, M. (1996). Refining the risk concept: Decision context as a factor mediating the relation between risk and program effectiveness. *Crime and Delinquency, 42*(3), 435-455.

Champion, D. J. (1994). *Measuring Offender Risk: A Criminal Justice Sourcebook.* London: Greenwood Press.

Coolican, H. (1994). *Research Methods and Statistics in Psychology.* London: Hodder and Stroughton.

Coulson, G., Ilacqua, G., Nutbrown, V., Giulekas, D., & Cudjoe, F. (1996). Predictive utility of the LSI for incarcerated female offenders. *Criminal Justice and Behavior, 23*(3), 427-439.

Criminal Code of Western Australia. (1983) Perth: Government Printer.

DeJong, J., Virkkunen, M., & Linnoila, M. (1992). Factors associated with recidivism in a criminal population. *The Journal of Nervous and Mental Disease, 180*(9), 543-550.

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.

Eggelston, E. (1976). *Fear, Favour or Affection: Aborigines and the law in Victoria, South Australia and Western Australia.* Canberra: Australian National University Press.

Elliot, D. S. (1994). Serious violent offenders: Onset, developmental course, and termination. *Criminology, 32*(1), 1-21.

Farrington, D. (1982). Longitudinal analysis of criminal violence. In M. E. Wolfgang & N. A. Weiner (Eds.), *Criminal Violence.* California: Sage Publications.

Feeley, M. M., & Simon, J. (1992). The new penology: Notes on the emerging strategy of corrections and its implications. *Criminology, 30*(4), 449-474.

Gardner, W., Lidz, C. W., Mulvey, E. P., & Shaw, E. C. (1996a). Clinical versus actuarial predictions of violence in patients with mental illness. *Journal of Consulting and Clinical Psychology, 64*(3), 602-609.

Gardner, W., Lidz, C. W., Mulvey, E. P., & Shaw, E. C. (1996b). A comparison of actuarial methods for identifying repetetively violent patients with mental illnesses. *Law and Human Behaviour, 20*(1), 35-48.

Gendreau, P., Little, T., & Goggin, C. (1996). A meta-analysis of the predictors of adult offender recidivism: What works! *Criminology, 34*(4), 575-607.

Gilmore, D. R., & Walkey, F. H. (1981). Identifying violent offenders using a measure of personal distance. *Journal of Consulting and Clinical Psychology, 49,* 287-291.

Grisso, T., & Appelbuam, P. S. (1992). Is it unethical to offer predictions of future violence? *Law and Human Behavior, 16*, 621-633.

Hanley, J. A., & McNeil, B. J. (1982). The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology, 143,* 29-36.

Harris, G. T., Rice, M. E., & Cormier, C. A. (1991). Psychopathy and violent recidivism. *Law and Human Behaviour, 15*(6), 625-637.

Harris, G. T., Rice, M. E., & Quinsey, V. L. (1993). Violent recidivism of mentally disordered offenders: The development of a statistical prediction instrument. *Criminal Justice and Behaviour, 20*(4), 315-335.

Harris, P. M. (1994). Client management classification and prediction probation outcome. *Crime and Delinquency, 40*(2), 154-174.

Hoffman, P. B. (1983). Screening for risk: A revised salient factor score (SFS81). *Journal of Criminal Justice, 11*, 539-47.

Howells, K. (1996). The psychological management of violence in clinical and forensic settings: Pitfalls and remedies. *Psychiatry, Psychology and Law, 3*, 71-76.

Howells, K., Watt, B., Hall, G., & Baldwin, S. (1997). Developing programmes for violent offenders. *Legal and Criminological Psychology, 2*, 117-128.

Indermaur, D. (1990). *Perceptions of Crime Seriousness and Sentencing.* Unpublished manuscript, Criminology Research Council, Western Australia.

Jones, P. R. (1993). Risk predictions in criminal justice.*The International Association of Residential and Community Alternatives' "What works in community corrections: A consensus conference"*. Philadelphia: Pennsylvania.

Katsiyannis, A., & Archwamety, T. (1997). Factors related to recidivism among delinquent youths in a state correctional facility. *Journal of Child and Family Studies, 6*(1), 43-55.

Klassen, D. & O'Connor, W. A. (1988a). A prospective study of predictors of violence in adult male mental health admissions. *Law and Human Behaviour, 12*(2), 143-158.

Klassen, D. & O'Connor, W. A. (1988b). Predicting violence in schizophrenic and non-schizophrenic patients: A prospective study. *Journal of Community Psychology, 16*, 217-227.

Klassen, D. & O'Connor, W. A. (1994). Demographic and case history variables in risk assessment. In J. Monahan and H. J. Steadman (Eds.), *Violence and*

Mental Disorder: Developments in Risk Assessment. Chicago: The University of Chicago Press.

Kraus, J. (1979). Trends in violent crime and public concern. *Australian Journal of Social Issues, 14*(3), 175-191.

Lincoln, R., & Wilson, P. (1994). Aboriginal Offending: Patterns and causes. In D. Chappell and P. Wilson (Eds.), *The Australian Criminal Justice System: The mid 1990s* (pp. 61-86). Australia: Butterworths.

Loeber, R., & Dishion, T. (1983). Early predictors of male delinquency: A review. *Psychological Bulletin, 94*(1), 68-99.

Loza, W., & Simourd, D. J. (1994). Psychometric evaluation of the Level of Supervision Inventory (LSI) among male Canadian federal offenders. *Criminal Justice and Behaviour, 21*(4), 468-480.

Martin, D. W. (1996). *Doing Psychology Experiments.* California: Brooks/Cole Publishing Company.

Martinez, R. Jr. (1997). Predictors of serious violent recidivism: Results from a cohort study. *Journal of Interpersonal Violence, 12*(2), 216-228.

McNiel, D. E., & Binder, R. L. (1994). Screening for risk of inpatient violence. *Law and Human Behavior, 18*(5), 579-593.

Megargee, E. I. (1981). Methodological problems in the prediction of violence. In J. R. Hays, T. K. Roberts, & K. S. Solway (Eds.), *Violence and the Violent Individual* (pp. 179-191). New York: Spectrum Publications.

Metz, C. E., Shen, J., Wang, P. & Kronman, H. E. (1989). ROCFIT [Computer software]. University of Chicago. Illinois: Authors.

Monahan, J. (1981).*Predicting Violent Behaviour: An Assessment of Clinical Techniques*. Beverly Hills CA: Sage Publications.

Monahan, J. (1984). The prediction of violent behavior: Toward a second generation of theory and policy. *American Journal of Psychiatry, 141*(1), 10-15.

Monahan, J. (1992). Mental disorder and violent behavior. *American Psychologist, 47*(4), 511-521.

Monahan, J. & Steadman, H. J. (1994). *Violence and Mental Disorder.* Chicago: Chicago University Press.

Morgan, R. (1994). Imprisonment. In M. Maguire, R. Morgan, & R. Reiner (Eds.), *The Oxford Handbook of Criminology.* Oxford: Clarendon Press.

Mossman, D. (1994a). Assessing predictions of violence: Being accurate about accuracy. *Journal of Consulting and Clinical Psychology, 62*(4), 783-792.

Mossman, D. (1994b). Further comments on portraying the accuracy of violence predictions. *Law and Human Behavior, 18*(5), 587-593.

Mossman, D., & Somoza, E. (1992). Balancing risks and other benefits: Another approach to optimizing diagnostic tests. *Journal of Neuropsychiatry, 4*(3), 331-335.

Motiuk, M. S., Motiuk, L. L., & Bonta, J. (1992). A comparison between self-report and interview-based inventories in offender classification. *Criminal Justice and Behavior, 19*(2), 143-159.

Motiuk, L. L., Bonta, J., & Andrews, J. A. (1986). Classification in correctional halfway houses: The relative and incremental predictive criterion validities of the Megargee-MMPI and LSI systems. *Criminal Justice and Behavior, 13*(1), 33-46.

Norusis, M. J. / SPSS Inc. (1990). SPSS/PC + advanced statistics 4.0 for the IBM PC/XT/AT and PS/2: SPSS statistical data analysis. Chicago: SPSS Inc.

Petersilia, J. (1978). The validity of criminality data derived from personal interviews. In C. Wellford (Ed.). *Quantitative Studies in Criminology.* Beverly Hills CA: Sage Publications.

Pratt, J. (1995). Dangerousness, risk and technologies of power. *The Australian and New Zealand Journal of Criminology, 28,* 3-31.

Prins, H. (1996). Risk assessment and management in criminal justice and psychiatry. *The Journal of Forensic Psychiatry, 7*(1), 42-62.

Ressler, R. K., Douglas, J. E., Burgess, A. W., & Burgess, A. G. (1992). *Crime Classification Manual: The Standard System for Investigating and Classifying Violent Crimes.* Great Britain: Simon & Schuster.

Rice, M. E., & Harris, G. T. (1992). A comparison of criminal recidivism among schizophrenic and nonschizophrenic offenders. *International Journal of Law and Psychiatry, 15,* 397-408.

Rice, M. E., & Harris, G. T. (1995a). Violent recidivism: Assessing predictive validity. *Journal of Consulting and Clinical Psychology, 63*(5), 737-748.

Rice, M. E., & Harris, G. T. (1995b). Psychopathy, schizophrenia, alcohol abuse, and violent recidivism. *International Journal of Law and Psychiatry,18*(3), 333-342.

Selby, M. J. (1984). Assessment of violence potential using measures of anger, hostility, and social desirability. *Journal of Personality Assessment, 48*(5), 531-544.

Shaffer, C. E. Jr., Waters, W. F., & Adams, S. G. Jr. (1994). Dangerousness: Assessing the risk of violent behavior. *Journal of Consulting and Clinical Psychology, 62*(5), 1064-1068.

Shah, S. A. (1981). Dangerousness: Conceptual, prediction and public policy issues. In J. R. Hays, T. K. Roberts, & K. S. Solway (Eds.), *Violence and the Violent Individual* (pp. 151-178). New York: Spectrum Publications.

Steadman, H. J., & Cocozza, J. J. (1974). *Careers of the Criminally Insane.* United States of America: D. C. Heath and Company.

Swets, J. (1986a). Indices of discrimination or diagnostic accuracy: Their ROCs and implied models. *Psychological Bulletin, 99*(1), 100-117.

Swets, J. (1986b). Form of empirical ROCs in discrimination and diagnostic tasks: Implications for theory and measurement of performance. *Psychological Bulletin, 99*(2), 181-198.

Thornberry, T. P., & Jacoby, J. E. (1979). *The Criminally Insane.* Chicago: The University of Chicago Press.

Virkkunen, M., DeJong, J., Bartko, J., Goodwin, F. K., & Linnoila, M., (1989). Relationship of psychobiological variables to recidivism in violent offenders and impulsive fire setters. *Archives of General Psychiatry, 46,* 600-603.

Walker, J. (1989). Prison sentences in Australia: Estimates of the characteristics of offenders sentenced to prison in 1987-88. *Trends and Issues: In Crime and Criminal Justice.* No. 20.

Ward, A. & Baldwin, S. (in press). Anger and violence management programs in correctional services: An annotated bibliography. *The Prison Journal.*

Webster, C. D., Harris, G. T., Rice, M. E., Cormier, C., Quinsey, V. L. (1994). *The Violence Prediction Scheme: Assessing Dangerousness in High Risk Men.* Canada: Centre of Criminology.

Wenk, E., Robison, J. D., & Smith, G. W. (1972). Can violence be predicted? *Crime and Delinquency, 18*, 393-402.

Appendices

Appendix A (i)

The Violent Offender Treatment Programs Risk Screening Assessment tool.

VOTP Risk Assessment Scale

		SCORE
1. CURRENT VIOLENT OFFENCE	Violence without bodily harm	(Score 1) ____
	Violence with bodily harm	(Score 2) ____
	Injuries life threatening	(Score 3) ____
	Injuries causing death	(Score 5) ____
2 MOST SERIOUS OFFENCE NOT INCLUDING CURRENT OFFENCE	Non violent	(Score 0) ____
	Violence without bodily harm (Score 1) ____	
	Violence with bodily harm (Score 2) ____	
	Injuries life threatening	(Score 3) ____
	Injuries causing death	(Score 4) ____
3. PREVIOUS VIOLENT OFFENCES	No previous convictions	(Score 0) ____
	1 previous conviction	(Score 2) ____
	2-3 previous convictions	(Score 4) ____
	More than 3 previous convictions	(Score 6) ____
4. PREVIOUS NON VIOLENT OFFENCES	No previous convictions	(Score 0) ____
	1 previous conviction	(Score 1) ____
	2-4 previous convictions	(Score 2) ____
	5 or more convictions	(Score 3) ____
5. AGE AT FIRST OFFENCE	Age 25 or more	(Score 1) ____
	Age 21-24	(Score 2) ____
	Age 15-20	(Score 3) ____
	Age 14 or below	(Score 4) ____
6. ALCOHOL RELATED OFFENCES	Non drinker of alcohol	(Score 0) ____
	Occasional use of alcohol	(Score 1) ____
	Binge drinker	(Score 2) ____
	Moderate regular use of alcohol (Score 3) ____	
	Heavy regular use of alcohol	(Score 4) ____
7. OTHER DRUG MISUSE	Non user of drugs	(Score 0) ____
	Occasional user, Non-intravenous	(Score 1) ____
	Moderate heavy user, Non-intravenous	(Score 2) ____
	Intravenous drug user	(Score 3) ____
	Poly Drug user	(Score 4) ____
		TOTAL SCORE ____

Appendix A (ii)

The Violent Offender Treatment Programme: Risk Screening Assessment Instruction Sheet.

For the purpose of the Risk Assessment Screening Assessment, violent offence is determined by reference to the criminal code.

1. Current Violent Offence
Examine the current violent offence for the extent of injury suffered by the victim. Write the corresponding score in the right hand column.

2. Most Serious Offence (Other Than Current)
Examine the offender's criminal record (both adult and juvenile). Determine which category the most serious previous offence falls into. Do not include current offence in this section. Write the score in the right hand column.

3. Previous Violent Offences
Use the offender's criminal history (both adult and juvenile) to determine how many violent offences have been recorded. DO NOT include offences committed at the same time as the current offence. Write the score in the right hand column.

4. Previous Non-Violent Offences
Use the offender's criminal history (both adult and juvenile) to determine how many non-violent offences have been recorded. DO NOT include offences committed at the same time as the current offence. Write the score in the right hand column.

5. Age at First Offence
Use the offender's criminal history (both adult and juvenile), social history and judge's comments, to determine the age of the offender when the first offence was committed. Write the score in the right hand column.

6. Use of Alcohol
Examine the offender's social history, judge's comments and any other available information to determine the offender's use of alcohol. If no use is recorded then consider the offender a "non-drinker", regardless of your personal opinion. Write the score in the right hand column.

7. Other Drug Use
Examine the offender's social history, judge's comments and any other available information to determine the offender's use of drugs other than alcohol. If no use is recorded then consider the offender a "non-user", regardless of your personal opinion.

8. Total Score
Add all scores from items 1 to 7 and write next to "Total Score".

Appendix B

The revised Violent Offender Treatment Program Risk Screening tool.

VOTP Risk Assessment Scale

		SCORE
(A) LEVEL OF HARM		
1. CURRENT VIOLENT OFFENCE	Violence without bodily harm	(Score 1) ____
	Violence with bodily harm	(Score 2) ____
	Grievous bodily harm	(Score 3) ____
	Injuries causing death	(Score 5) ____
2 MOST SERIOUS OFFENCE NOT INCLUDING CURRENT OFFENCE	Non violent	(Score 0) ____
	Violence without bodily harm	(Score 1) ____
	Violence with bodily harm	(Score 2) ____
	Grievous bodily harm	(Score 3) ____
	Injuries causing death	(Score 4) ____
(B) PROBABILITY		
3. PREVIOUS VIOLENT OFFENCES	No previous convictions	(Score 0) ____
	1 previous conviction	(Score 2) ____
	2-3 previous convictions	(Score 4) ____
	More than previous convictions	(Score 6) ____
4. PREVIOUS NON VIOLENT OFFENCES ____	No previous convictions	(Score 0)
	1 previous conviction	(Score 1) ____
	2-4 previous convictions	(Score 2) ____
	5 or more convictions	(Score 3) ____
5. AGE AT FIRST OFFENCE	Age 25 or more	(Score 1) ____
	Age 21-24	(Score 2) ____
	Age 15-20	(Score 3) ____
	Age 14 or below	(Score 4) ____
6. ALCOHOL RELATED OFFENCES	Non user of alcohol	(Score 0) ____
	1 alcohol offence	(Score 1) ____
	2 alcohol offences	(Score 2) ____
	3 alcohol offences	(Score 3) ____
	4 or more alcohol offences	(Score 4) ____
7. OTHER DRUG MISUSE	Non user of drugs	(Score 0) ____
	Cannabis	(Score 1) ____
	Other Non-intravenous drugs	(Score 2) ____
	Intravenous drug user	(Score 3) ____
	Poly Drug user	(Score 4) ____
		TOTAL SCORE ____

Appendix C

The Subsequent Offence Form.

Subsequent Offence Form

		SCORE 1 FOR EACH OFFENCE
1. SUBSEQUENT OFFENCE TYPE	Non violent	_____
	Violence without bodily harm	_____
	Violence with bodily harm	_____
	Grievous bodily harm	_____
	Injuries causing death	_____

2. TIME BETWEEN RELEASE AND FIRST FAILURE

Years_____ Months_____

3. TIME BETWEEN RELEASE AND FIRST VIOLENT FAILURE

Years_____ Months_____

Appendix D

Groupings of the Criminal Code of Western Australia (1983).

Violence without bodily harm

S. 68	Going armed so as to cause fear.
S. 71	Affray.
S. 72	Challenge to a fight or duel.
S. 74	Threatening violence.
S. 75	Interfering with political liberty.
S. 78	Punishment of piracy.
S. 79	Attempted piracy with personal violence.
S. 123	Corrupting or threatening jurors.
S. 128	Threatening witness before Royal Commission.
S. 179	Offering violence to officiating ministers of religion.
S. 180	Disturbing religious worship.
S. 223	Assaults, unlawful.
S. 284	Accessory after the fact to murder.
S. 285	Written threats to murder.
S. 286	Conspiring to murder.
S. 294A	Dangerous goods on aircrafts.
S. 295	Preventing escape from wreck.
S. 296	Intentionally endangering safety of persons travelling by railway.
S. 296A	Intentionally endangering the safety of persons travelling by aircraft.
S. 298	Causing explosion likely to endanger life.
S. 299	Attempting to cause explosion likely to endanger life.
S. 300	Maliciously administering poison with intent to harm.
S. 305	Setting mantraps.
S. 307	Endangering safety of persons travelling by railroad.
S. 308	Sending or taking unseaworthy ships to sea.
S. 309	Endangering steamships by tampering with machinery.
S. 310	The like by engineers.
S. 312	Landing explosives.
S. 313	Common assault.
S. 314	Assault with intent to commit unnatural offence.
S. 316	Assaults on persons protecting wrecks.
S. 318	Serious assaults.
S. 318A	Assaults on members of crew or aircraft.
S. 321	Common assaults.
S. 322	Aggravated assaults.
S. 324	Assaults in interference with freedom of trade or work.
S. 329	Abduction.
S. 330	Abduction of girls under sixteen.
S. 332	Kidnapping.
S. 333	Deprivation of liberty.

S. 334	False certificates or other documents by officers charged with duties relating to liberty.
S. 335	Concealment of matters affecting liberty.
S. 336	Procuring the apprehension or detention of a person not suffering from mental disorder.
S. 337	Unlawful custody of persons suffering mental disorder.
S. 338	Threats.
S. 391	Definition of robbery.
S. 392	Loaded arms.
S. 393	Punishment of robbery.
S. 394	Attempted robbery accompanied by wounding, or in company. (In company only)
S. 395	Assaults with intent to steal.
S. 397	Demanding property by written threats.
S. 398	Attempts at extortion by threats.
S. 399	Procuring execution of deeds, etc., by threats.
S. 463A	Threats to safety of aircraft.
S. 463B	False statements relating to aircraft.
S. 550	Intimidation or annoyance by violence or otherwise.

Violence with bodily harm

S. 292	Disabling in order to commit indictable offence.
S. 293	Stupefying in order to commit indictable offence.
S. 294	Acts intended to cause grevious bodily harm or prevent arrest.
S. 301	Wounding and similar acts.
S. 302	Failure to supply necessaries.
S. 303	Endangering life or health of apprentices or servants.
S. 304	Endangering life of children by exposure.
S. 306	Negligent acts causing harm.
S. 315	Indecent assault on males.
S. 317	Assaults occasioning bodily harm.
S. 324A	Assaults occasioning bodily harm.
S. 325	Definition of rape.
S. 326	Punishment of rape.
S. 327	Attempt to commit rape.
S. 328	Indecent assaults on females.
S. 394	Attempted robbery accompanied by wounding, or in company. (Wounding only).

Grevious bodily harm

S. 283	Attempt to murder.
S. 297	Grevious bodily harm.

Injuries causing death

S. 261	Consent to death immaterial.

S. 268	Killing of a human being unlawful.
S. 271	Deaths by acts done at childbirth.
S. 272	Causing death by threats.
S. 273	Acceleration of death.
S. 274	When injury or death might be prevented by proper precaution.
S. 275	Injuries causing death in consequence of subsequent treatment.
S. 277	Unlawful homicide.
S. 278	Definition of wilful murder.
S. 279	Definition of murder.
S. 280	Definition of manslaughter.
S. 282	Punishment of wilful murder and murder.
S. 287	Punishment of manslaughter.
S. 288	Aiding suicide.
S. 290	Killing unborn child.

Appendix E (i)

Data collapsed into categories for the 10-year time-at-risk interval. Groupings are indicated by solid lines.

Score	Nonviolent Recidivist	Violent Recidivist
3	4	0
4	4	1
5	8	1
6	9	1
7	13	3
8	14	5
9	9	8
10	5	4
11	8	5
12	6	8
13	1	7
14	6	3
15	6	6
16	6	12
17	1	7
18	1	3
19	1	7
20	2	5
21	2	5
22	0	2
23	1	2
Total	107	95

Appendix E (ii)

Data collapsed into categories for the 36month time-at-risk interval. Categories are indicated by solid lines.

Score	Nonviolent Recidivist	Violent Recidivist
3	3	1
4	4	1
5	9	0
6	10	0
7	13	3
8	16	3
9	13	4
10	7	2
11	11	2
12	8	6
13	6	3
14	8	1
15	8	4
16	8	10
17	5	3
18	1	3
19	4	4
20	2	5
21	2	4
22	1	2
23	1	1
Total	143	59

Appendix E (iii)

Data collapsed into categories for the 60 month time-at-risk interval. Categories are indicated by solid lines.

Score	Nonviolent Recidivists	Violent Recidivists
3	4	0
4	4	1
5	8	1
6	10	0
7	13	3
8	16	3
9	12	5
10	7	2
11	9	4
12	7	7
13	3	6
14	7	2
15	7	5
16	7	11
17	3	5
18	1	3
19	4	4
20	2	5
21	2	4
22	1	2
23	1	1
Total	128	74

Appendix E (iv)

Data collapsed into working categories for the 84 month time-at-risk interval. Categories are indicated by solid lines.

Score	Nonviolent Recidivists	Violent Recidivists
3	4	0
4	4	1
5	8	1
6	10	0
7	13	3
8	14	5
9	10	7
10	7	2
11	9	4
12	6	8
13	1	8
14	6	3
15	6	6
16	7	11
17	2	6
18	1	3
19	2	6
20	2	5
21	2	4
22	1	2
23	1	1
Total	116	86

Appendix F (i)

Calculated operating points for the 10 year time-at-risk interval.

Cutoff Points	Sensitivity	1 - Specificity
22	.02	.01
21	.04	.01
20	.09	.03
19	.15	.05
18	.22	.06
17	.25	.07
16	.33	.08
15	.45	.13
14	.52	.19
13	.56	.24
12	.62	.25
11	.70	.31
10	.76	.38
9	.80	.43
8	.88	.51
7	.94	.65
6	.97	.77
5	.98	.85
4	.99	.93
3	1	.97

Appendix F (ii)

Calculated operating points for the 36 month time-at-risk interval.

Cutoff Point	Sensitivity	1 - Specificity
22	.01	.01
21	.05	.01
20	.11	.03
19	.19	.04
18	.26	.07
17	.31	.08
16	.35	.11
15	.52	.17
14	.58	.23
13	.60	.29
12	.64	.33
11	.74	.39
10	.77	.47
9	.81	.52
8	.87	.61
7	.92	.72
6	.97	.82
5	.97	.89
4	.97	.95
3	.98	.98

Appendix F (iii)

Calculated operating points for the 60 month time-at-risk interval.

Cutoff Point	Sensitivity	1 - Specificity
22	.01	.01
21	.04	.02
20	.09	.03
19	.16	.05
18	.22	.08
17	.26	.09
16	.32	.11
15	.47	.16
14	.54	.22
13	.56	.27
12	.70	.30
11	.78	.35
10	.80	.42
9	.82	.48
8	.89	.57
7	.93	.70
6	.97	.80
5	.97	.88
4	.99	.94
3	1	.97

Appendix F (iv)

Calculated operating points for the 84 month time-at-risk interval.

Cutoff Point	Sensitivity	1 - Specificity
22	.01	.01
21	.03	.02
20	.08	.04
19	.14	.05
18	.21	.07
17	.24	.08
16	.31	.10
15	.44	.16
14	.51	.21
13	.55	.26
12	.64	.27
11	.73	.32
10	.78	.40
9	.80	.46
8	.88	.54
7	.94	.66
6	.98	.78
5	.98	.86
4	.99	.93
3	1	.97

Appendix G (i)

Two by two contingency Tables showing classification for each cutoff point for the 10-year time-at-risk interval.

Cutoff 3

(TP) 95	(FP) 0
(FN) 103	(TN) 4

Cutoff 4

(TP) 94	(FP) 1
(FN) 99	(TN) 8

Cutoff 5

(TP) 93	(FP) 2
(FN) 91	(TN) 16

Cutoff 6

(TP) 92	(FP) 3
(FN) 82	(TN) 25

Cutoff 7

(TP) 89	(FP) 6
(FN) 69	(TN) 38

Cutoff 8

(TP) 84	(FP) 11
(FN) 55	(TN) 52

Cutoff 9

(TP) 76	(FP) 19
(FN) 46	(TN) 61

Cutoff 10

(TP) 72	(FP) 23
(FN) 41	(TN) 66

Cutoff 11

(TP) 67	(FP) 28
(FN) 33	(TN) 74

Cutoff 12

(TP) 59	(FP) 36
(FN) 27	(TN) 80

Cutoff 13

(TP) 52	(FP) 43
(FN) 26	(TN) 81

Cutoff 14

(TP) 49	(FP) 46
(FN) 20	(TN) 87

Cutoff 15

(TP) 43	(FP) 52
(FN) 14	(TN) 93

Cutoff 16

(TP) 31	(FP) 64
(FN) 8	(TN) 99

Cutoff 17

(TP) 24	(FP) 71
(FN) 7	(TN) 100

Cutoff 18

(TP) 21	(FP) 74
(FN) 6	(TN) 101

Cutoff 19

(TP) 14	(FP) 81
(FN) 5	(TN) 102

Cutoff 20

(TP) 9	(FP) 86
(FN) 3	(TN) 104

Cutoff 21

(TP) 4	(FP) 91
(FN) 1	(TN) 106

Cutoff 22

(TP) 2	(FP) 93
(FN) 1	(TN) 106

Cutoff 23

(TP) 0	(FP) 95
(FN) 0	(TN) 107

Appendix G (ii)

Two by two contingency Tables showing classification for each cutoff point for the 36 month time-at-risk interval.

Cutoff 3

(TP) 61	(FP) 1
(FN) 137	(TN) 3

Cutoff 4

(TP) 60	(FP) 2
(FN) 133	(TN) 7

Cutoff 5

(TP) 60	(FP) 2
(FN) 124	(TN) 16

Cutoff 6

(TP) 60	(FP) 2
(FN) 114	(TN) 26

Cutoff 7

(TP) 57	(FP) 5
(FN) 101	(TN) 39

Cutoff 8

(TP) 54	(FP) 8
(FN) 85	(TN) 55

Cutoff 9

(TP) 50	(FP) 12
(FN) 72	(TN)68

Cutoff 10

(TP) 48	(FP) 14
(FN) 65	(TN) 75

Cutoff 11

(TP) 46	(FP) 16
(FN) 54	(TN) 86

Cutoff 12

(TP) 40	(FP) 22
(FN) 46	(TN) 94

Cutoff 13

(TP) 37	(FP) 25
(FN) 40	(TN) 100

Cutoff 14

(TP) 36	(FP) 26
(FN) 32	(TN) 108

Cutoff 15

(TP) 32	(FP) 30
(FN) 24	(TN) 116

Cutoff 16

(TP) 22	(FP) 40
(FN) 16	(TN) 124

Cutoff 17

(TP) 19	(FP) 43
(FN) 11	(TN) 129

Cutoff 18

(TP) 16	(FP) 46
(FN) 10	(TN) 130

Cutoff 19

(TP) 12	(FP) 50
(FN) 6	(TN) 134

Cutoff 20

(TP) 7	(FP) 55
(FN) 4	(TN) 136

Cutoff 21

(TP) 3	(FP) 59
(FN) 2	(TN) 138

Cutoff 22

(TP) 1	(FP) 61
(FN) 1	(TN) 139

Cutoff 23

(TP) 0	(FP) 62
(FN) 0	(TN) 140

Appendix G (iii)

Two by two contingency Tables showing classification for each cutoff point for the 60 month time-at-risk interval.

Cutoff 3		
	(TP) 74	(FP) 0
	(FN) 124	(TN) 4

Cutoff 4		
	(TP) 73	(FP) 1
	(FN) 120	(TN) 8

Cutoff 5		
	(TP) 72	(FP) 2
	(FN) 112	(TN) 16

Cutoff 6		
	(TP) 72	(FP) 2
	(FN) 102	(TN) 26

Cutoff 7		
	(TP) 69	(FP) 5
	(FN) 89	(TN) 39

Cutoff 8		
	(TP) 66	(FP) 8
	(FN) 73	(TN) 55

Cutoff 9		
	(TP) 61	(FP) 13
	(FN) 61	(TN) 67

Cutoff 10		
	(TP) 59	(FP) 15
	(FN) 54	(TN) 74

Cutoff 11		
	(TP) 55	(FP) 19
	(FN) 45	(TN) 83

Cutoff 12		
	(TP) 52	(FP) 26
	(FN) 38	(TN) 90

Cutoff 13		
	(TP) 42	(FP) 32
	(FN) 35	(TN) 93

Cutoff 14		
	(TP) 40	(FP) 34
	(FN) 28	(TN) 100

Cutoff 15		
	(TP) 35	(FP) 39
	(FN) 21	(TN) 107

Cutoff 16		
	(TP) 24	(FP) 50
	(FN) 14	(TN) 114

Cutoff 17		
	(TP) 19	(FP) 55
	(FN) 11	(TN) 117

Cutoff 18		
	(TP) 16	(FP) 58
	(FN) 10	(TN) 118

Cutoff 19		
	(TP) 12	(FP) 62
	(FN) 6	(TN) 122

Cutoff 20		
	(TP) 7	(FP) 67
	(FN) 4	(TN) 124

Cutoff 21		
	(TP) 3	(FP) 71
	(FN) 2	(TN) 126

Cutoff 22		
	(TP) 1	(FP) 73
	(FN) 1	(TN) 127

Cutoff 23		
	(TP) 0	(FP) 74
	(FN) 0	(TN) 128

Appendix G (iv)

Two by two contingency Tables showing the classification for each cutoff point for the 84 month time-at-risk interval.

Cutoff 3

(TP) 86	(FP) 0
(FN) 112	(TN) 4

Cutoff 4

(TP) 85	(FP) 1
(FN) 108	(TN) 8

Cutoff 5

(TP) 84	(FP) 2
(FN) 100	(TN) 16

Cutoff 6

(TP) 84	(FP) 2
(FN) 90	(TN) 26

Cutoff 7

(TP) 81	(FP) 5
(FN) 77	(TN) 39

Cutoff 8

(TP) 76	(FP) 10
(FN) 63	(TN) 53

Cutoff 9

(TP) 69	(FP) 17
(FN) 53	(TN) 63

Cutoff 10

(TP) 67	(FP) 19
(FN) 46	(TN) 70

Cutoff 11

(TP) 63	(FP) 23
(FN) 37	(TN) 79

Cutoff 12

(TP) 55	(FP) 31
(FN) 31	(TN) 85

Cutoff 13

(TP) 47	(FP) 39
(FN) 30	(TN) 86

Cutoff 14

(TP) 44	(FP) 42
(FN) 24	(TN) 92

Cutoff 15

(TP) 38	(FP) 48
(FN) 18	(TN) 98

Cutoff 16

(TP) 27	(FP) 59
(FN) 11	(TN) 105

Cutoff 17

(TP) 21	(FP) 65
(FN) 9	(TN) 107

Cutoff 18

(TP) 18	(FP) 68
(FN) 8	(TN) 108

Cutoff 19

(TP) 12	(FP) 74
(FN) 6	(TN) 110

Cutoff 20

(TP) 7	(FP) 79
(FN) 4	(TN) 112

Cutoff 21

(TP) 3	(FP) 83
(FN) 2	(TN) 114

Cutoff 22

(TP) 1	(FP) 85
(FN) 1	(TN) 115

Cutoff 23

(TP) 0	(FP) 86
(FN) 0	(TN) 116

Appendix H (i)

ROCFIT output for the 10-year time-at-risk period.

R O C F I T (IBM VERSION 1.2) :

M A X I M U M L I K E L I H O O D E S T I M A T I O N
O F A B I N O R M A L R O C C U R V E
F R O M R A T I N G D A T A

DATA DESCRIPTION: 10 year

DATA COLLECTED IN 9 CATEGORIES
WITH CATEGORY 9 REPRESENTING STRONGEST EVIDENCE OF POSITIVITY
(E.G., THAT ABNORMALITY IS PRESENT).

NO. OF ACTUALLY NEGATIVE CASES = 107.
NO. OF ACTUALLY POSITIVE CASES = 95.

RESPONSE DATA:

CATEGORY	1	2	3	4	5	6	7	8	9
ACTUALLY NEGATIVE CASES	38.	14.	14.	8.	6.	7.	6.	9.	5.
ACTUALLY POSITIVE CASES	6.	5.	12.	5.	8.	10.	6.	29.	14.

OBSERVED OPERATING POINTS:
FPF: .000 .047 .131 .187 .252 .308 .383 .514 .645 1.000
TPF: .000 .147 .453 .516 .621 .705 .758 .884 .937 1.000

INITIAL VALUES OF PARAMETERS:
A= 1.1220 B= 1.2284
Z(K)= -.371 -.035 .297 .500 .667 .889 1.122 1.678
LOGL= -402.9125
GOODNESS OF FIT --
CHI SQUARE= 7.0971 WITH 6 DEGREES OF FREEDOM, P= .3120

PROCEDURE CONVERGES AFTER 4 ITERATIONS.

FINAL VALUES OF PARAMETERS:

A= 1.1154 B= 1.2200

Z(K)= -.364 -.047 .316 .483 .662 .882 1.049 1.750

LOGL= -400.6775

GOODNESS OF FIT --

CHI SQUARE= 2.8738 WITH 6 DEGREES OF FREEDOM, P= .8245

VARIANCE-COVARIANCE MATRIX:

A	.0375	.0156	.0143	.0145	.0140	.0135	.0128	.0117	.0107	.0053
B	.0156	.0288	.0066	.0047	.0014	-.0006	-.0030	-.0062	-.0089	-.0216
Z(1)	.0143	.0066	.0151	.0119	.0095	.0086	.0077	.0067	.0059	.0027
Z(2)	.0145	.0047	.0119	.0137	.0110	.0101	.0092	.0083	.0076	.0052
Z(3)	.0140	.0014	.0095	.0110	.0132	.0122	.0114	.0107	.0103	.0090
Z(4)	.0135	-.0006	.0086	.0101	.0122	.0135	.0127	.0121	.0117	.0111
Z(5)	.0128	-.0030	.0077	.0092	.0114	.0127	.0143	.0137	.0135	.0137
Z(6)	.0117	-.0062	.0067	.0083	.0107	.0121	.0137	.0161	.0161	.0172
Z(7)	.0107	-.0089	.0059	.0076	.0103	.0117	.0135	.0161	.0183	.0203
Z(8)	.0053	-.0216	.0027	.0052	.0090	.0111	.0137	.0172	.0203	.0382

CORRELATION MATRIX:

A	1.0000	.4742	.6029	.6407	.6310	.6016	.5531	.4741	.4064	.1398
B	.4742	1.0000	.3168	.2372	.0698	-.0310	-.1467	-.2883	-.3857	-.6498
Z(1)	.6029	.3168	1.0000	.8319	.6754	.6040	.5254	.4268	.3545	.1129
Z(2)	.6407	.2372	.8319	1.0000	.8200	.7432	.6607	.5581	.4827	.2255
Z(3)	.6310	.0698	.6754	.8200	1.0000	.9190	.8357	.7345	.6604	.4005
Z(4)	.6016	-.0310	.6040	.7432	.9190	1.0000	.9171	.8183	.7463	.4909
Z(5)	.5531	-.1467	.5254	.6607	.8357	.9171	1.0000	.9039	.8344	.5864
Z(6)	.4741	-.2883	.4268	.5581	.7345	.8183	.9039	1.0000	.9331	.6945
Z(7)	.4064	-.3857	.3545	.4827	.6604	.7463	.8344	.9331	1.0000	.7650
Z(8)	.1398	-.6498	.1129	.2255	.4005	.4909	.5864	.6945	.7650	1.0000

AREA = .7603 STD. DEV.(AREA) = .0336

ESTIMATED BINORMAL ROC CURVE, WITH LOWER AND UPPER

BOUNDS ON ASYMMETRIC 95% CONFIDENCE INTERVAL FOR
TRUE-POSITIVE FRACTION AT EACH SPECIFIED
FALSE-POSITIVE FRACTION:

FPF	TPF	(LOWER BOUND, UPPER BOUND)
.005	.0213	(.0027 , .1016)
.010	.0424	(.0081 , .1488)
.020	.0822	(.0230 , .2158)
.030	.1191	(.0412 , .2668)
.040	.1537	(.0613 , .3093)
.050	.1863	(.0825 , .3463)
.060	.2172	(.1045 , .3794)
.070	.2466	(.1268 , .4095)
.080	.2746	(.1492 , .4371)
.090	.3014	(.1717 , .4628)
.100	.3270	(.1940 , .4868)
.110	.3516	(.2161 , .5094)
.120	.3752	(.2379 , .5307)
.130	.3979	(.2593 , .5509)
.140	.4197	(.2804 , .5701)
.150	.4408	(.3010 , .5884)
.200	.5354	(.3975 , .6692)
.250	.6152	(.4824 , .7356)
.300	.6830	(.5565 , .7911)
.400	.7901	(.6770 , .8758)
.500	.8677	(.7691 , .9325)
.600	.9228	(.8409 , .9678)
.700	.9603	(.8976 , .9875)
.800	.9839	(.9422 , .9966)
.900	.9963	(.9765 , .9996)
.950	.9991	(.9899 , 1.0000)

ESTIMATES OF EXPECTED OPERATING POINTS ON FITTED ROC CURVE, WITH LOWER AND UPPER BOUNDS OF ASYMMETRIC 95% CONFIDENCE INTERVAL ALONG THE CURVE FOR THOSE POINTS:

EXPECTED OPERATING POINT (FPF, TPF)	LOWER BOUND (FPF, TPF)	UPPER BOUND (FPF, TPF)
(.0400, .1539)	(.0164, .0684)	(.0858, .2903)
(.1472, .4349)	(.0944, .3128)	(.2167, .5635)
(.1888, .5155)	(.1289, .3956)	(.2633, .6341)
(.2540, .6210)	(.1852, .5090)	(.3344, .7236)
(.3145, .7006)	(.2387, .5983)	(.3991, .7892)
(.3762, .7675)	(.2945, .6759)	(.4639, .8425)
(.5188, .8796)	(.4278, .8142)	(.6088, .9268)
(.6421, .9406)	(.5492, .8973)	(.7273, .9681)

Appendix H (ii)

ROCFIT output for the 36 month time-at-risk- interval.

R O C F I T (IBM VERSION 1.2) :

M A X I M U M L I K E L I H O O D E S T I M A T I O N
O F A B I N O R M A L R O C C U R V E
F R O M R A T I N G D A T A

DATA DESCRIPTION: 36 month

DATA COLLECTED IN 7 CATEGORIES
WITH CATEGORY 7 REPRESENTING STRONGEST EVIDENCE OF POSITIVITY
(E.G., THAT ABNORMALITY IS PRESENT).

NO. OF ACTUALLY NEGATIVE CASES = 140.
NO. OF ACTUALLY POSITIVE CASES = 62.

RESPONSE DATA:

CATEGORY	1	2	3	4	5	6	7
ACTUALLY NEGATIVE CASES	39.	29.	26.	22.	13.	5.	6.
ACTUALLY POSITIVE CASES	5.	7.	10.	8.	13.	7.	12.

OBSERVED OPERATING POINTS:
FPF: .000 .043 .079 .171 .329 .514 .721 1.000
TPF: .000 .194 .306 .516 .645 .806 .919 1.000

INITIAL VALUES OF PARAMETERS:
A= .8420 B= .9593
Z(K)= -.587 -.036 .443 .948 1.415 1.719
LOGL= -365.3356
GOODNESS OF FIT --
CHI SQUARE= 2.1581 WITH 4 DEGREES OF FREEDOM, P= .7067

PROCEDURE CONVERGES AFTER 4 ITERATIONS.

FINAL VALUES OF PARAMETERS:

A= .8446 B= .9586

Z(K)= -.586 -.033 .456 .905 1.417 1.759

LOGL= -364.8975

GOODNESS OF FIT --

CHI SQUARE= 1.3348 WITH 4 DEGREES OF FREEDOM, P= .8554

VARIANCE-COVARIANCE MATRIX:

A	.0325	.0114	.0081	.0081	.0078	.0071	.0056	.0042
B	.0114	.0194	.0039	.0022	-.0002	-.0034	-.0084	-.0126
Z(1)	.0081	.0039	.0125	.0079	.0057	.0043	.0028	.0019
Z(2)	.0081	.0022	.0079	.0106	.0078	.0062	.0050	.0043
Z(3)	.0078	-.0002	.0057	.0078	.0109	.0091	.0080	.0076
Z(4)	.0071	-.0034	.0043	.0062	.0091	.0135	.0124	.0123
Z(5)	.0056	-.0084	.0028	.0050	.0080	.0124	.0212	.0213
Z(6)	.0042	-.0126	.0019	.0043	.0076	.0123	.0213	.0312

CORRELATION MATRIX:

A	1.0000	.4521	.4023	.4375	.4137	.3385	.2141	.1308
B	.4521	1.0000	.2509	.1521	-.0153	-.2106	-.4156	-.5123
Z(1)	.4023	.2509	1.0000	.6880	.4894	.3302	.1752	.0956
Z(2)	.4375	.1521	.6880	1.0000	.7245	.5212	.3317	.2345
Z(3)	.4137	-.0153	.4894	.7245	1.0000	.7473	.5238	.4106
Z(4)	.3385	-.2106	.3302	.5212	.7473	1.0000	.7326	.6004
Z(5)	.2141	-.4156	.1752	.3317	.5238	.7326	1.0000	.8282
Z(6)	.1308	-.5123	.0956	.2345	.4106	.6004	.8282	1.0000

AREA = .7290 STD. DEV.(AREA) = .0389

ESTIMATED BINORMAL ROC CURVE, WITH LOWER AND UPPER BOUNDS ON ASYMMETRIC 95% CONFIDENCE INTERVAL FOR TRUE-POSITIVE FRACTION AT EACH SPECIFIED FALSE-POSITIVE FRACTION:

FPF	TPF	(LOWER BOUND, UPPER BOUND)
.005	.0521	(.0121 , .1593)
.010	.0829	(.0252 , .2073)
.020	.1304	(.0511 , .2693)
.030	.1688	(.0760 , .3138)
.040	.2021	(.1000 , .3497)
.050	.2319	(.1230 , .3803)
.060	.2591	(.1451 , .4073)
.070	.2842	(.1664 , .4316)
.080	.3076	(.1870 , .4538)
.090	.3297	(.2068 , .4744)
.100	.3505	(.2260 , .4936)
.110	.3702	(.2445 , .5117)
.120	.3890	(.2624 , .5288)
.130	.4070	(.2798 , .5450)
.140	.4242	(.2966 , .5604)
.150	.4408	(.3130 , .5752)
.200	.5151	(.3878 , .6410)
.250	.5786	(.4530 , .6966)
.300	.6339	(.5105 , .7448)
.400	.7265	(.6079 , .8239)
.500	.8008	(.6883 , .8846)
.600	.8615	(.7572 , .9301)
.700	.9110	(.8185 , .9628)
.800	.9507	(.8750 , .9843)
.900	.9809	(.9298 , .9962)
.950	.9923	(.9587 , .9991)

ESTIMATES OF EXPECTED OPERATING POINTS ON FITTED ROC CURVE, WITH LOWER AND UPPER BOUNDS OF ASYMMETRIC 95% CONFIDENCE INTERVAL ALONG THE CURVE FOR THOSE POINTS:

EXPECTED OPERATING POINT (FPF , TPF)	LOWER BOUND (FPF , TPF)	UPPER BOUND (FPF , TPF)
(.0393, .2001)	(.0176, .1203)	(.0789, .3053)
(.0782, .3036)	(.0443, .2155)	(.1288, .4050)
(.1828, .4909)	(.1287, .4048)	(.2491, .5774)
(.3242, .6582)	(.2543, .5836)	(.4009, .7271)
(.5131, .8095)	(.4329, .7526)	(.5927, .8576)
(.7210, .9202)	(.6431, .8842)	(.7895, .9469)

Appendix H (iii)

ROCFIT output for the 60 month time-at-risk interval.

```
      R O C F I T (IBM VERSION 1.2) :

  M A X I M U M  L I K E L I H O O D  E S T I M A T I O N
     O F  A  B I N O R M A L  R O C  C U R V E
        F R O M  R A T I N G  D A T A

DATA DESCRIPTION: 60month

DATA COLLECTED IN 8 CATEGORIES
WITH CATEGORY 8 REPRESENTING STRONGEST EVIDENCE OF POSITIVITY
(E.G., THAT ABNORMALITY IS PRESENT).

NO. OF ACTUALLY NEGATIVE CASES = 128.
NO. OF ACTUALLY POSITIVE CASES = 74.

RESPONSE DATA:
CATEGORY                1   2   3   4   5   6  7   8
ACTUALLY NEGATIVE CASES 39. 28. 16. 7. 17. 10. 5. 6.
ACTUALLY POSITIVE CASES 5. 8. 6. 7. 13. 16. 7. 12.

OBSERVED OPERATING POINTS:
FPF: .000 .047 .086 .164 .297 .352 .477 .695 1.000
TPF: .000 .162 .257 .473 .649 .743 .824 .932 1.000

INITIAL VALUES OF PARAMETERS:
A= .9899  B= 1.1546
Z(K)= -.511 .059 .381 .533 .978 1.366 1.676
LOGL= -386.9041
GOODNESS OF FIT --
CHI SQUARE= 3.5623 WITH 5 DEGREES OF FREEDOM, P= .6140

PROCEDURE CONVERGES AFTER 4 ITERATIONS.
```

FINAL VALUES OF PARAMETERS:

A= .9846 B= 1.1427

Z(K)= -.500 .050 .346 .531 .951 1.408 1.707

LOGL= -386.1428

GOODNESS OF FIT --

CHI SQUARE= 1.8580 WITH 5 DEGREES OF FREEDOM, P= .8684

VARIANCE-COVARIANCE MATRIX:

A	.0352	.0140	.0107	.0110	.0107	.0103	.0090	.0067	.0049
B	.0140	.0249	.0051	.0028	.0007	-.0009	-.0054	-.0114	-.0158
Z(1)	.0107	.0051	.0132	.0086	.0071	.0063	.0048	.0032	.0022
Z(2)	.0110	.0028	.0086	.0114	.0095	.0086	.0072	.0059	.0052
Z(3)	.0107	.0007	.0071	.0095	.0114	.0104	.0090	.0080	.0076
Z(4)	.0103	-.0009	.0063	.0086	.0104	.0118	.0104	.0096	.0094
Z(5)	.0090	-.0054	.0048	.0072	.0090	.0104	.0147	.0144	.0147
Z(6)	.0067	-.0114	.0032	.0059	.0080	.0096	.0144	.0224	.0230
Z(7)	.0049	-.0158	.0022	.0052	.0076	.0094	.0147	.0230	.0312

CORRELATION MATRIX:

A	1.0000	.4713	.4961	.5501	.5344	.5055	.3936	.2395	.1474
B	.4713	1.0000	.2818	.1653	.0409	-.0540	-.2826	-.4845	-.5681
Z(1)	.4961	.2818	1.0000	.7042	.5822	.5084	.3447	.1891	.1104
Z(2)	.5501	.1653	.7042	1.0000	.8375	.7461	.5529	.3720	.2782
Z(3)	.5344	.0409	.5822	.8375	1.0000	.8983	.6923	.5033	.4041
Z(4)	.5055	-.0540	.5084	.7461	.8983	1.0000	.7853	5921	.4903
Z(5)	.3936	-.2826	.3447	.5529	.6923	.7853	1.0000	.7922	.6836
Z(6)	.2395	-.4845	.1891	.3720	.5033	.5921	.7922	1.0000	.8707
Z(7)	.1474	-.5681	.1104	.2782	.4041	.4903	.6836		

AREA = .7416 STD. DEV.(AREA) = .0353

ESTIMATED BINORMAL ROC CURVE, WITH LOWER AND UPPER BOUNDS ON ASYMMETRIC 95% CONFIDENCE INTERVAL FOR TRUE-POSITIVE FRACTION AT EACH SPECIFIED FALSE-POSITIVE FRACTION:

FPF	TPF	(LOWER BOUND, UPPER BOUND)
.005	.0250	(.0039 , .1044)
.010	.0470	(.0105 , .1494)
.020	.0865	(.0270 , .2123)
.030	.1220	(.0458 , .2599)
.040	.1547	(.0659 , .2995)
.050	.1853	(.0866 , .3340)
.060	.2140	(.1076 , .3649)
.070	.2413	(.1286 , .3930)
.080	.2672	(.1497 , .4189)
.090	.2919	(.1705 , .4430)
.100	.3156	(.1911 , .4656)
.110	.3383	(.2114 , .4869)
.120	.3601	(.2314 , .5071)
.130	.3811	(.2510 , .5263)
.140	.4013	(.2702 , .5446)
.150	.4208	(.2891 , .5622)
.200	.5092	(.3772 , .6402)
.250	.5848	(.4553 , .7056)
.300	.6502	(.5243 , .7613)
.400	.7566	(.6396 , .8494)
.500	.8376	(.7313 , .9119)
.600	.8986	(.8059 , .9539)
.700	.9433	(.8679 , .9798)
.800	.9742	(.9199 , .9936)
.900	.9928	(.9636 , .9990)
.950	.9979	(.9825 , .9999)

ESTIMATES OF EXPECTED OPERATING POINTS ON FITTED ROC CURVE, WITH LOWER AND UPPER BOUNDS OF ASYMMETRIC 95% CONFIDENCE INTERVAL ALONG THE CURVE FOR THOSE POINTS:

EXPECTED OPERATING POINT (FPF , TPF)	LOWER BOUND (FPF , TPF)	UPPER BOUND (FPF , TPF)
(.0439, .1670)	(.0200, .0866)	(.0868, .2842)
(.0795, .2660)	(.0444, .1685)	(.1324, .3861)
(.1708, .4592)	(.1172, .3541)	(.2379, .5673)
(.2977, .6472)	(.2284, .5533)	(.3752, .7328)
(.3647, .7222)	(.2895, .6370)	(.4456, .7962)
(.4799, .8230)	(.3977, .7544)	(.5630, .8781)
(.6915, .9402)	(.6085, .9031)	(.7658, .9651)

Appendix H (iv)

ROCFIT output for the 84 month time-at-risk interval.

R O C F I T (IBM VERSION 1.2):

M A X I M U M L I K E L I H O O D E S T I M A T I O N
O F A B I N O R M A L R O C C U R V E
F R O M R A T I N G D A T A

DATA DESCRIPTION: 84 month

DATA COLLECTED IN 10 CATEGORIES
WITH CATEGORY 10 REPRESENTING STRONGEST EVIDENCE OF POSITIVITY
(E.G., THAT ABNORMALITY IS PRESENT).

NO. OF ACTUALLY NEGATIVE CASES = 116.
NO. OF ACTUALLY POSITIVE CASES = 86.

RESPONSE DATA:

CATEGORY	1	2	3	4	5	6	7	8	9	10
ACTUALLY NEGATIVE CASES	39.	14.	10.	16.	6.	7.	6.	7.	5.	6.
ACTUALLY POSITIVE CASES	5.	5.	7.	6.	8.	11.	17.	15.	5.	7.

OBSERVED OPERATING POINTS:
FPF: .000 .052 .095 .155 .207 .267 .319 .457 .543 .664 1.000
TPF: .000 .081 .140 .314 .512 .640 .733 .802 .884 .942 1.000

INITIAL VALUES OF PARAMETERS:
A= 1.0854 B= 1.4907
Z(K)= -.422 -.108 .108 .470 .621 .817 1.014 1.312 1.629
LOGL= -433.1142
GOODNESS OF FIT --
CHI SQUARE= 13.8856 WITH 7 DEGREES OF FREEDOM, P= .0533
PROCEDURE CONVERGES AFTER 4 ITERATIONS.

FINAL VALUES OF PARAMETERS:

A= 1.0742 B= 1.4697

Z(K)= -.403 -.106 .116 .379 .543 .754 1.047 1.413 1.667

LOGL= -429.1634

GOODNESS OF FIT --

CHI SQUARE= 6.4895 WITH 7 DEGREES OF FREEDOM, P= .4839

VARIANCE-COVARIANCE MATRIX:

A	.0439	.0183	.0156	.0161	.0160	.0153	.0146	.0135	.0116	.0089	.0068
B	.0183	.0374	.0074	.0057	.0037	.0004	-.0019	-.0053	-.0103	-.0167	-.0210
Z(1)	.0156	.0074	.0140	.0112	.0097	.0083	.0075	.0065	.0052	.0037	.0027
Z(2)	.0161	.0057	.0112	.0127	.0110	.0095	.0087	.0078	.0067	.0054	.0046
Z(3)	.0160	.0037	.0097	.0110	.0121	.0106	.0098	.0090	.0081	.0071	.0064
Z(4)	.0153	.0004	.0083	.0095	.0106	.0121	.0114	.0107	.0100	.0094	.0091
Z(5)	.0146	-.0019	.0075	.0087	.0098	.0114	.0125	.0119	.0114	.0111	.0110
Z(6)	.0135	-.0053	.0065	.0078	.0090	.0107	.0119	.0137	.0134	.0135	.0137
Z(7)	.0116	-.0103	.0052	.0067	.0081	.0100	.0114	.0134	.0170	.0175	.0181
Z(8)	.0089	-.0167	.0037	.0054	.0071	.0094	.0111	.0135	.0175	.0244	.0251
Z(9)	.0068	-.0210	.0027	.0046	.0064	.0091	.0110				

CORRELATION MATRIX:

A	1.0000	.4528	.6284	.6809	.6919	.6651	.6247	.5502	.4244	.2706	.1814
B	.4528	1.0000	.3245	.2622	.1713	.0199	-.0896	-.2326	-.4064	-.5528	-.6052
Z(1)	.6284	.3245	1.0000	.8388	.7430	.6341	.5631	.4680	.3386	.2014	.1279
Z(2)	.6809	.2622	.8388	1.0000	.8866	.7666	.6915	.5918	.4555	.3078	.2261
Z(3)	.6919	.1713	.7430	.8866	1.0000	.8726	.7965	.6972	.5610	.4106	.3246
Z(4)	.6651	.0199	.6341	.7666	.8726	1.0000	.9239	.8279	.6970	.5489	.4603
Z(5)	.6247	-.0896	.5631	.6915	.7965	.9239	1.0000	.9062	.7801	.6356	.5468
Z(6)	.5502	-.2326	.4680	.5918	.6972	.8279	.9062	1.0000	.8785	.7389	.6505
Z(7)	.4244	-.4064	.3386	.4555	.5610	.6970	.7801	.8785	1.0000	.8598	.7701
Z(8)	.2706	-.5528	.2014	.3078	.4106	.5489	.6356	.7389	.8598	1.0000	.8951
Z(9)	.1814	-.6052	.1279	.2261	.3246	.4603	.5468	.6505	.7701	.8951	1.0000

AREA = .7272 STD. DEV.(AREA) = .0349

ESTIMATED BINORMAL ROC CURVE, WITH LOWER AND UPPER BOUNDS ON ASYMMETRIC 95% CONFIDENCE INTERVAL FOR TRUE-POSITIVE FRACTION AT EACH SPECIFIED FALSE-POSITIVE FRACTION:

FPF	TPF	(LOWER BOUND, UPPER BOUND)
.005	.0033	(.0002 , .0328)
.010	.0095	(.0009 , .0595)
.020	.0259	(.0041 , .1060)
.030	.0455	(.0098 , .1471)
.040	.0669	(.0178 , .1848)
.050	.0895	(.0278 , .2197)
.060	.1129	(.0395 , .2526)
.070	.1367	(.0528 , .2836)
.080	.1608	(.0674 , .3132)
.090	.1850	(.0831 , .3414)
.100	.2091	(.0998 , .3684)
.110	.2331	(.1172 , .3943)
.120	.2569	(.1353 , .4192)
.130	.2805	(.1539 , .4432)
.140	.3038	(.1729 , .4663)
.150	.3267	(.1922 , .4887)
.200	.4355	(.2904 , .5898)
.250	.5332	(.3861 , .6758)
.300	.6195	(.4752 , .7487)
.400	.7588	(.6276 , .8598)
.500	.8586	(.7465 , .9312)
.600	.9259	(.8374 , .9718)
.700	.9674	(.9055 , .9912)
.800	.9896	(.9542 , .9983)
.900	.9985	(.9858 , .9999)
.950	.9998	(.9954 , 1.0000)

ESTIMATES OF EXPECTED OPERATING POINTS ON FITTED ROC CURVE, WITH LOWER AND UPPER BOUNDS OF ASYMMETRIC 95% CONFIDENCE INTERVAL ALONG THE CURVE FOR THOSE POINTS:

EXPECTED OPERATING POINT (FPF, TPF)	LOWER BOUND (FPF, TPF)	UPPER BOUND (FPF, TPF)
(.0477, .0844)	(.0217, .0291)	(.0942, .1953)
(.0789, .1582)	(.0428, .0733)	(.1342, .2904)
(.1476, .3211)	(.0963, .2003)	(.2144, .4647)
(.2254, .4863)	(.1626, .3550)	(.3000, .6192)
(.2936, .6088)	(.2230, .4817)	(.3730, .7251)
(.3523, .6975)	(.2761, .5795)	(.4350, .7978)
(.4540, .8171)	(.3701, .7214)	(.5400, .8891)
(.5423, .8907)	(.4543, .8174)	(.6283, .9401)
(.6564, .9522)	(.5677, .9074)	(.7372, .9776)

Appendix I

Computer output for inter-rater reliability, Cohen's Kappa.

RATER1 by RATER2

Count	RATER2 0	1	2	3	4	Row Total
RATER1 0	39	1				40
						29.2
1	1	22	4		1	28
						20.4
2	1		11	1	1	14
						10.2
3		1	2	27		31
						22.6
4	1			1	15	17
						12.4
5						3
						2.2
6						4
						2.9
Column Total	42	24	17	29	17	137
	30.7	17.5	12.4	21.2	12.4	100.0

(Continued)

RATER1 by RATER2

RATER1	Count	RATER2 5	RATER2 6	Row Total
0				40
				29.2
1				28
				20.4
2				14
				10.2
3		1		31
				22.6
4				17
				12.4
5		3		3
				2.2
6			4	4
				2.9
Column Total		4 2.9	4 2.9	137 100.0

Statistic	Value	ASE1	Val/ASE0	Approximate Significance
Kappa	.85351	.03413	20.63340	.00000

Number of Missing Observations: 3

RATER1 by RATER3

RATER1	Count	RATER3 0	1	2	3	4	Row Total
	0	39	1				40
							29.2
	1	2	25			1	28
							20.4
	2	1	1	9	1	1	14
							10.2
	3	1	1	3	25		31
							22.6
	4	2			2	13	17
							12.4
	5						3
							2.2
	6					1	4
							2.9
	Column	45	28	12	28	16	137
(Continued)	Total	32.8	20.4	8.8	20.4	11.7	100.0

RATER1 by RATER3

RATER1	Count	RATER3 5	RATER3 6	Row Total
0				40
				29.2
1				28
				20.4
2			1	14
				10.2
3		1		31
				22.6
4				17
				12.4
5		3		3
				2.2
6			3	4
				2.9
Column Total		4 2.9	4 2.9	137 100.0

Statistic	Value	ASE1	Val/ASE0	Approximate Significance
Kappa	.81548	.03766	19.45158	.00000

Number of Missing Observations: 3

RATER2 by RATER3

RATER2	Count	RATER3 0	1	2	3	4	Row Total
0		44					44
							31.4
1		1	22	1			24
							17.1
2			7	9		1	18
							12.9
3		1		2	24	1	29
							20.7
4		1			3	13	17
							12.1
5					1		4
							2.9
6						1	4
							2.9
	Column	47	29	12	28	16	140
(Continued)	Total	33.6	20.7	8.6	20.0	11.4	100.0

RATER2 by RATER3

Count	RATER3 5	RATER3 6	Row Total
RATER2 0			44
			31.4
RATER2 1			24
			17.1
RATER2 2		1	18
			12.9
RATER2 3	1		29
			20.7
RATER2 4			17
			12.1
RATER2 5	3		4
			2.9
RATER2 6		3	4
			2.9
Column Total	4	4	140
	2.9	2.9	100.0

Statistic	Value	ASE1	Val/ASE0	Approximate Significance
Kappa	.80134	.03793	19.49345	.00000

Number of Missing Observations: 0

Appendix J

Computer output of Spearman's rho correlation coefficients for item-by-item and item-by-total.

```
- - -  S P E A R M A N   C O R R E L A T I O N   C O E F F I C I E N T S  - - -

ITEM2          .2292
            N(  202)
            Sig .001

ITEM3          .2301       .9051
            N(  202)    N(  202)
            Sig .001    Sig .000

ITEM4          .0093       .3620       .4071
            N(  202)    N(  202)    N(  202)
            Sig .896    Sig .000    Sig .000

ITEM5         -.0248       .1790       .1855       .4072
            N(  202)    N(  202)    N(  202)    N(  202)
            Sig .726    Sig .011    Sig .008    Sig .000

ITEM6          .1836       .3998       .4829       .3075       .0263
            N(  202)    N(  202)    N(  202)    N(  202)    N(  202)
            Sig .009    Sig .000    Sig .000    Sig .000    Sig .710

ITEM7         -.2087      -.2269      -.2333      -.0307       .0058      -.2521
            N(  202)    N(  202)    N(  202)    N(  202)    N(  202)    N(  202)
            Sig .003    Sig .001    Sig .001    Sig .664    Sig .935    Sig .000

TOTAL          .3387       .8103       .8671       .6268       .3668       .6212
            N(  202)    N(  202)    N(  202)    N(  202)    N(  202)    N(  202)
            Sig .000    Sig .000    Sig .000    Sig .000    Sig .000    Sig .000

               ITEM1       ITEM2       ITEM3       ITEM4       ITEM5       ITEM6


TOTAL         -.0905
            N(  202)
            Sig .200

               ITEM7
```

(Coefficient / (Cases) / 2-tailed Significance)

" . " is printed if a coefficient cannot be computed