The Prediction Of Risk Level And Violent Offending Using Models From Classical And Bounded Rationality

Grant Broad

Submitted for the degree of
Doctor of Clinical Psychology
(D. Clin. Psych.)

Psychology Department

City University

London

February 2004

Contents – Volume One

Acknowledgements

I would like to express my gratitude to the following people for the help they have given me during this project:

My wife Vanessa for her support and understanding when the pressures of job and research made me question my commitment to doctorate study.

My supervisor, Dr Peter Ayton, for helping me identify a theoretical framework of direct relevance both to my employment and the types of questions I wanted this project to answer.

Mrs Jill Brennan and Mrs Helen Rootes, for endless patience in the face of huge rough drafts, and without whom none of this would have appeared in its finished state. They can carve on their office doors with pride, "She never lost her sense of humour".

Mr Christopher Tomlin of the Kent Police for access to information about those who had offended.

Dr Mandeep Dhami and Ms Sarah Keen for help with data analysis

Kent Forensic Psychiatry Service managers Dr Lona Roberts and Mr Kevin Halpin. Despite the time and financial pressures which impact on how our directorate functions, they have always facilitated my efforts to complete this study.

Copyright

Section A: Preface

Introduction To The Portfolio

What first drew me to the study of psychology was the question, "Why do people do what they do?" After 23 years of practising as a clinical psychologist, the last 15 of them involving forensic clients, I have come to the realisation that although there are general conclusions which can be reached, in most instances there are as many answers to this question as there are people to ask it of. It has been difficult to accept that despite clear evidence of the potential for trouble from certain types of behaviour, people continue to engage in it, often to the detriment of everyone involved. Sexual offenders continue to break the law despite having been punished, those who know that intoxication is a precursor to their violence keep drinking or taking illicit substances, and despite the absence of any sign of positive interest from a selected individual, stalkers persist with efforts to intrude on their victim's privacy.

With each particular moral and legal transgression there are fascinating threads to unravel in an attempt to help the person learn how their acts have come about. Of just as much interest to me are the processes people go through when making decisions about what to do, or not to do, next. How does someone decide to go down one pathway as opposed to another? This question is far from trivial. Every person, no matter what their circumstances, faces a multitude of decision points each day, and learning more about how judgements are made has a real urgency

about it which has justified to me, my employers, and others in my life taking time over the last three years to investigate the process more closely.

I began my doctoral study with a critical review of the psychological literature on the controversy between clinical and actuarial approaches to decision making (Section D). In the former, evidence may be weighed and evaluated subjectively to make a prediction, while the actuarial method involves collecting and interpreting data in a mathematical way, where the probabilities of particular acts occurring are calculated on the basis of population statistics. Repeated comparisons since the 1950s suggest that actuarial judgements are more likely to be accurate than ones based on accumulated clinical wisdom alone (Grove and Meehl, 1996). Some areas investigated have been diagnosis from MMPI scores, the presence or absence of brain dysfunction, who should receive places on university courses, parole violation, response to ECT, interpretation of pathology results, determining delinquency from EEG records, and violent re-offending (Meehl, 1954). The last area is of particular interest to me on a number of levels. As a scientist-practitioner I want to refine my judgements about who is most at risk of committing violence so that resources are not inappropriately allocated, the clients' civil liberties are infringed minimally if at all, and people in the community are protected from the worst consequences of foreseeable risk.

One would think that practitioners would welcome methods to enhance the accuracy of their decisions, but this is not always so. Failure to accept a consistent body of scientific evidence over unvalidated personal observation may be seen as a normal human shortcoming, but in the case of professionals who identify themselves as

working from empirical foundations it might also be viewed as plainly irrational (Dawes et al, 1989). Resistance is based on beliefs such as the unique nature of the practitioner's work setting, the problem of generalising predictions from aggregates to the individual, qualitative data may be thought to have more bearing on clinical decisions than numerical material, the patient is an individual not an object, the more experience a clinician has the better their predictions are, statistical decision rules will quickly become out of date because the world is constantly changing, and many others. Challenges to these arguments are presented in the review, with the conclusion being that although actuarial approaches to decision making do not have all the answers, they are at least transparent, explicit, allow for informed criticism, and are readily available to other scientists wishing to replicate or extend research.

After completing the literature review I felt better equipped to turn to the main part of the portfolio (Section B). My research topic was the application of reasoning models to decisions about and predictions of violent offending amongst patients referred to the Kent Forensic Psychiatry Service for risk assessment. The models compared were logistic regression and a comparatively new way of looking at how judgements are made called the matching heuristic. The former has been used in the majority of psychological research about inductive inference in clinical judgement (Brehmer, 1994). The assumption is that before a decision is reached all the relevant information will be considered, assigned its proper weights, and combined together in some way to produce the best answer. This view has been challenged by those who believe that human inference is systematically biased and error prone, and that the principles governing decision making are simple rules rather that the elaborate computations stemming from probability theory (Kahneman et al, 1982). One of the

simplest is the matching heuristic, where only one piece of information is used to reach a decision. In this study, decision targets were accurate division of the patient cohort into high risk, low risk, those who had offended violently, or those who had not offended violently by time of follow up.

One hundred psychiatrists were contacted through postal survey, and asked for the cues they look for when assessing the potential a patient has to commit violence. The answers provided by the 37 respondents were placed in rank order, and the ten most popular (which also received support from the empirical literature on risk prediction) were used as a basis for data collection. One hundred patient risk assessment reports done by Kent Forensic Psychiatry Service staff were checked for the presence or absence of these cues in a retrospective file study. Note was taken of the report writer's risk conclusion at the time (low or high). A search was conducted on the Police National Computer to determine if the patients concerned had subsequently offended violently (OV) or not (NOV).

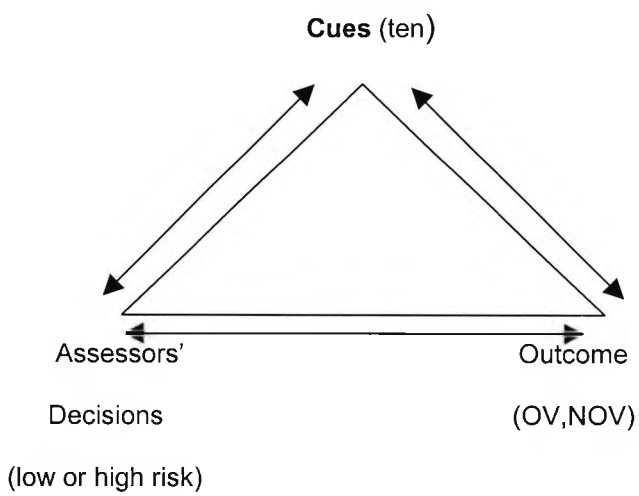
The following comparisons were possible;

Figure 1 INTERACTIONS BETWEEN STUDY VARIABLES

Adapted from

Figure 13.2, page 238,

In Jackson (1997)



There were 12 research hypotheses. They concerned the accuracy with which the three models used could separate the 100 patients in the study cohort into the group which offended violently, the group which did not offend violently, those who were assessed as at high risk of violent offending, and those who were rated as being at low risk of committing violence. The 12 null hypotheses were that none of the three models used would be able to separate patients accurately into the offended violently, not offended violently, high risk, or low risk groups any better than chance.

Table Fourteen below shows the accuracy with which the models separated patients in relation to outcome and assessors decisions;

TABLE FOURTEEN : ACCURACY OF MODELS IN SEPARATING PATIENTS WITH REGARD TO OUTCOME AND ASSESSORS' DECISIONS

		OUTCOME		ASSESSORS' DECISIONS	
		Offended Violently Group correct separation	Not Offended Violently Group correct separation	High Risk Group correct separation	Low Risk Group correct separation
MODELS USED	Logistic Regression (10 cues)	0%	88%	39%	21%
	Logistic Regression (two most significant cues)	0%	88%	39%	21%
	Matching Heuristic (one cue) with the highest level of accuracy (ie. threat/ control override permitting phenomena)	0%	72%	12%	50%

Table Sixteen below illustrates the research hypotheses, and whether or not they were supported by the study data;

TABLE SIXTEEN: RESEARCH HYPOTHESES AND FINDINGS

		OUTCOME		ASSESSORS' DECISIONS	
		Patients who offended violently	Patients who did not Offend violently	Patients assessed as high risk of committing violence	Patients assessed as low risk of committing violence
MODELS USED	Logistic Regression (10 cues)	ONE accurate separation at a better than chance rate Not supported	TWO accurate separation at a better than chance rate Supported	THREE accurate separation at a better than chance rate Not supported	FOUR accurate separation at a better than chance rate Not supported
	Logistic Regression (two most significant cues)	FIVE accurate separation at a better than chance rate Not supported	SIX accurate separation at a better than chance rate Supported	SEVEN accurate separation at a better than chance rate Not supported	EIGHT accurate separation at a better than chance rate Not supported
	Matching Heuristic using the cue with the highest level of predictive accuracy	NINE accurate separation at a better than chance rate Not supported	TEN accurate separation at a better than chance rate Supported	ELEVEN accurate separation at a better than chance rate Not supported	TWELVE accurate separation at a better than chance rate Not supported

The final part of the portfolio was a case study from my clinical practice, highlighting the dynamic tensions between judgement of risk, protection of people in the community, and consideration of the rights of the patient (Section C). I described the assessment, treatment, management, and follow up of an individual (alias Mr Field) convicted under the anti-stalking laws, and sentenced to imprisonment. His preoccupation with a media personality had intruded on her personal freedom and led to the curtailment of his own.

The concept of client-centred therapy (Rogers, 1951) and its core belief that practitioners must demonstrate unconditional positive regard (Truax, 1963, 1966) have shaped therapeutic interventions in a humanistic way, but when dealing with a forensic client one must also consider how personal freedoms interact with the rights of others. The challenge is to balance the needs of the client and those who live around them. My Theoretical approach to this piece of work was based on Rational-Emotive Therapy (Ellis, 1962), now known as Rational-Emotive Behavioural Counselling. There were 16 sessions over a 15-month period, and Mr Field was seen while detained on remand, during his prison sentence, and finally while living in the community under supervision by the Probation Service.

Working with Mr Field was my first professional exposure to stalking behaviour. I was surprised to discover that forensic psychiatry has given scant attention to this phenomenon (Kamphuis and Emmelkamp, 2000), despite the major impact it has on people. Hall (1998) reported that victims may become more cautious, suspicious, anxious and aggressive. Pathé and Mullen (1997) found the recipients of such unwanted attention might adjust their daily routines, obtain an unlisted telephone number, buy home security systems, change jobs, suffer decreased productivity at work, or even move to a new neighbourhood. It is clear that stalking behaviour can be much more than harmless and eccentric infatuation. At the beginning of our work together, Mr Field believed that the official reaction to what he had done was out of all proportion to the risk he posed. Because he saw his victim as a love object, someone he would never harm, it was difficult for him to appreciate how she might respond to the ambiguity present in some of his letters to her. There was also little appreciation of the social and political context of his behaviour. A media personality,

Jill Dando had been murdered in mysterious circumstances during the time that Mr Field was writing to his victim, making her fears for her safety more intense. Some of his attempts to reassure her were seen by the woman, and by the sentencing judge, as ambiguously menacing. As our sessions progressed it became more difficult for Mr Field to support his irrational beliefs about the offending behaviour, and he gradually moved away from a utilitarian perspective to a more genuinely therapeutic one. Our work together took place over a long enough time period for him to consolidate his gains as he went, and not to be overwhelmed by the speed of confrontation and change. At the time of writing (i.e. September 2003) Mr Field has not come to the attention of the police for any violent offending.

I do not view myself as an expert on stalking, but take heart from the proof this case has given me that a scientist-practitioner with a generic training base can apply first principles to new types of referrals, and make a contribution to understanding, treating, and managing them.

Section A References

1. Brehmer,B., (1994), The psychology of linear judgement models, Acta Psychologica, Vol.87, pp.137-154.

2. Dawes, R.M., Faust, D., and Meehl, P.E., (1989), Clinical versus actuarial judgement, Science, Vol. 243, pp.1668-1674.

3. Ellis, A., (1962), Reason and emotion in psychotherapy. Secaucus, New Jersey: Citadel

4. Grove, W.M., and Meehl, P.E., (1996), Comparative efficacy of informal (subjective, impressionistic) and formal (mechanical, algorithmic) prediction procedures: The clinical – statistical controversy, Psychology, Public Policy and Law, Vol. 2, No. 2, pp. 293-323.

5. Hall, D.M., (1998), The victims of stalking. In The psychology of stalking: Clinical and forensic perspectives (ed. J.R. Meloy), pp. 113-137, San Diego, CA: Academic Press

6. Kahneman, D., Slovic, P., and Tversky, A., (Eds), (1982) Judgement under uncertainty: Heuristics and biases. Cambridge, England: Cambridge University Press.

7. Kamphuis, J.H., and Emmelkamp, P.M.G., (2000), Stalking – a contemporary challenge for forensic and clinical psychiatry, British Journal of Psychiatry, Vol. 176, pp. 206-209.

8. Meehl, P.E., (1954), Clinical versus statistical prediction, a theoretical analysis and a review of the evidence, University of Minnesota Press, Minneapolis, MN.

9. Pathé, M., and Mullen, P.E., (1997), The impact of stalkers on their victims, British Journal of Psychiatry, Vol. 170, pp. 12-17.

10. Rogers, C.R., (1951), Client – centred therapy; its current practice. Boston: Houghton.

11. Truax, C.B., (1963), Effective ingredients in psychotherapy: An approach to unravelling the patient – therapist interaction, Journal of Counselling Psychology, Vol. 10, pp. 256-263

12. Truax, C.B., (1966), Therapist empathy, warmth, and genuineness and patient personality change in group psychotherapy: A comparison between interaction unit measures, time sample measures, and patient perception measures, Journal of Clinical Psychology, Vol. 22, pp. 225-229

The Prediction Of Risk Level And Violent Offending Using Models From Classical And Bounded Rationality

Abstract for Section B: Research

Cues Consultant Psychiatrists use to estimate the probability of future violence were obtained through postal survey. The ten most popular were taken as a basis for data collection. One hundred patient risk assessment reports done by Kent Forensic Psychiatry Service staff were checked in a retrospective file study for these cues. Note was also made of the report writer's risk conclusion at the time (i.e. low or high), and a search carried out on the Police National Computer to determine if the patients concerned had subsequently offended violently or not. There were 46% assessed as high risk, and 54% considered low risk. Over the follow up period (29 – 76 months) 12 had offended violently while 88 had not. Assessors' decisions could be predicted at a better than chance rate by using one cue in a non-compensatory model (the matching heuristic) or two cues in a compensatory model such as the logistic regression approach. These two showed little change in accuracy by comparison with a logistic regression model using all ten cues (i.e. 62%, 60% and 60% respectively). There was greater variability when attempting to forecast outcome (i.e. 72%, 88% and 88% respectively). Implications for clinical practice will be discussed.

Introduction To The Dissertation

When there are decisions to be made about important issues the automatic assumption is that all the relevant information will be considered, assigned its proper weights, and combined together in some way to produce the best answer. This premise has led to the development of statistical methods in an attempt to make real-world chaos orderly and less frightening. Multiple regression is both the economist's tool (McCloskey, 1985) and a model of inductive inference in clinical judgement (Brehmer, 1994). The Enlightenment view that probability theory and human reasoning are two sides of the same coin faded in the early 19th Century but has remained strong in psychology and economics (Gigerenzer and Goldstein, 1996). However, this viewpoint has been challenged by those who believe that human inference is systematically biased and error prone, suggesting that the principles governing decision making are simple rules rather than the elaborate computations stemming from probability theory (Kahneman, Slovic, and Tversky, 1982).

Applying the classical approach to the real world makes the mind appear to be a super-calculator, containing vast amounts of information which only needs time and concentration to produce a result. By contrast, the simpler view leads us to believe that people are hopelessly lost in the face of environmental complexity, and the often erratic nature of the decisions they make reflects this.

But there is a third way to look at inference, focussing on the psychological and ecological rather than logic and probability theory (Gigerenzer and Goldstein, 1996).

Herbert Simon has suggested that it might be more profitable to look for models of bounded rather than classical rationality (Simon, 1956, 1982). He argued that information processing systems typically need to satisfice rather than optimise. Satisficing is a word of Scottish origin. It is a blend of sufficing and satisfying which Simon uses to characterise heuristics (rules or set procedures for aiding the discovery of truth) which successfully deal with conditions of limited time, knowledge, or computational capacities. The assumption is that a decision-maker chooses the first option that satisfies their aspiration level rather than taking the time to survey all possible alternatives, estimating probabilities and utilities for the various outcomes associated with each alternative, calculating expected worth, and choosing the highest scorer.

Bounded rationality has both cognitive and ecological sides. Minds are adapted to real world environments: the two go in tandem (Simon, 1956); "Human rational behaviour is shaped by a scissors whose two blades are the structure of task environments and the computational capabilities of the actor", (Simon, 1990, page 7).

The insight that the minds of living systems should be understood relative to the environment in which they evolved has had little impact so far in research on human inference. Simple psychological heuristics have often been discredited without a fair trial because they looked "stupid by the norms of classical rationality", (Gigerenzer

and Goldstein, 1996, page 651). The authors conducted a test between rational heuristics and a satisficing one, using two alternative choice tasks where the person must make an inference based solely on knowledge retrieved from memory. Subjects were asked to use a list of cues to decide which one of a set of paired German cities had more inhabitants. The fast and frugal model outperformed all competitors in terms of inferential speed and accuracy, thus providing proof that cognitive mechanisms capable of successful performance do not need to satisfy classical norms of rational deduction.

Such findings have provided the impetus for this dissertation. The author wished to discover if a fast and frugal reasoning model could be applied to the prediction of risk level and violent offending. Clinicians asked to make risk assessments about patients are often under pressure from high case loads, limited time, and incomplete information. Any system which allows them to complete their task accurately while avoiding redundant activity could have significant implications for patient care, public safety and job satisfaction.

A series of cues consultant psychiatrists use to estimate the probability of future violence when assessing a patient were obtained through postal survey (see Chapter 3, Section 3.4). These were placed in rank order, and the ten most popular (which also received support from the empirical literature on risk prediction) were used as a basis for data collection. These may be seen in Appendix B2. One hundred patient risk assessment reports done by Kent Forensic Psychiatry Service staff (KFPS) were checked for the presence or absence of these cues. Note was taken of the report writer's risk conclusion at the time (i.e. Low or High), and a search

carried out on the Police National Computer to determine if the patients concerned had subsequently offended violently (OV) or not (NOV).

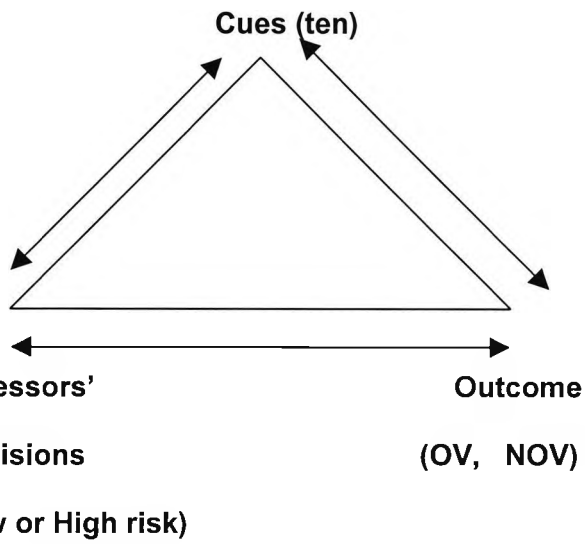
The following comparisons were possible;

Figure 1 **INTERACTIONS BETWEEN STUDY VARIABLES**

Adapted from

Figure 13.2, page 238,

in Jackson (1997)



There were 12 research hypotheses. They concerned the accuracy with which the three models used could separate the 100 patients in the study cohort into the group which offended violently, the group which did not offend violently, those who were assessed as at high risk of violent offending, and those who were rated as being at low risk of committing violence. The 12 null hypotheses were that none of the three models used would be able to separate patients accurately into the offended violently, not offended violently, high risk, or low risk groups any better than chance.

There are major implications for clinical practice. The way forensic mental health services conduct risk assessments at the moment involves a great deal of time, reliance on other people and agencies to provide what is deemed adequate background information to reach a conclusion, and anxiety. Practitioners worry about overlooking something crucial to the final decision, and as the stakes involved in making an error are high, this contributes to stress in the work place. If a handful of readily accessible items prove just as valid and reliable in predicting violent behaviour as the very labour intensive and costly process currently in place, then forensic services can reduce work load without sacrificing efficiency. On the other hand, if fast and frugal reasoning using satisficing heuristics does not make an equivalent or improved contribution to risk assessment, then protracted clinical interviews and collateral data gathering will continue to be the norm, but at least practitioners can console themselves that there is as yet no better way to arrive at a decision.

The reader will have to wait until later in this report to find the answer, but in preparation I would now like to turn to a consideration of the research literature on the prediction of violent offending.

Chapter 1

The Prediction of Violence

1.1 Definition of violence

Given that violence is at the core of risk assessment and many clinical decisions, it is surprising that its definition has received little attention in the research literature (Douglas, Cox and Webster, 1999). These authors adopt a view that is supported by others (e.g. Boer, Hart, Kropp and Webster 1998; Hart, 1998; Webster et al, 1997b). They see violence as actual, attempted, or threatened harm to a person or persons. This includes acts that can be reasonably considered to harm others. Threats ought to be unambiguous to count as violent. Boer et al (1998) consider sexual violence to be "actual, attempted, or threatened sexual contact with a person who is not consenting or unable to give consent" (page 9).

In the large scale McArthur risk assessment project in the United States, researchers created categories of violence depending on seriousness (Steadman et al, 1998). Violence was defined as "battery that resulted in physical injury, sexual assaults, assaultative acts that involved the use of a weapon, or threats made with a weapon in hand" (page 395). They then define "other aggressive acts" as "battery that did not result in physical injury" (page 395).

Zillman (1978) has classified violence as physical assault causing, or having the intention of causing, serious physical injury or death to others.

Blackburn (1995) has described it as the forceful and intended infliction of physical injury on an unwilling victim.

Although there is some agreement among researchers about what constitutes violence, the term is used imprecisely by many to cover behaviours as diverse as vandalism, joyriding, or youngsters jeering at the police (Blackburn, 1995).

What is agreed is that violence is part of the wider phenomenon of aggression, seen as the intended infliction of, or threats of inflicting, bodily injury upon another (Berkowitz, 1974; Zillman, 1978). Vigorous or destructive actions are carried out for many reasons, so the term aggression should be limited to behaviour in which the intended goal is harm or injury of another living being who is motivated to avoid it (Zillman, 1978; Berkowitz, 1993). This definition covers psychological as well as physical injury, and distinguishes malevolent from benevolently intended damage, such as surgical intervention. It also allows for separate consideration of harm which is accidental or sought masochistically by the victim.

Aggression, then, may entail physical attack, verbal abuse, or passive obstruction, and its consequences range from loss of life to psychological hurt (Blackburn, 1995). It may involve a specific act or might be used to describe a personal predisposition. A trait of aggressiveness implies a tendency to react with aggression under certain conditions, but not all offenders who have engaged in a violent crime have aggressive attitudes.

In summary, violence may be seen as part of the wider concept of aggression. It involves actual, attempted or threatened harm and is directed at another living being who is motivated to avoid it. The consequences may range from psychological distress to loss of life. It is such an important phenomenon, with the potential to impinge on most people's lives at some point, that it is no surprise researchers have been making efforts for many years to predict violence.

1.2 The history of violence prediction

Before 1966 relatively little attention was paid to how well clinicians assessed risk (Dolan and Doyle, 2000). The Baxstrom v Herald (1966) ruling in the United States resulted in release or transfer from maximum security hospitals of 966 patients to lower security settings or into the community. This cohort was followed up by Steadman and Cocozza (1974), who reported that after four years of liberty only 20% had been re-convicted, the majority for non-violent offences. Throughout the 1970's several other studies supported the notion that clinicians had little expertise in predicting violent outcomes (e.g. Cocozza and Steadman, 1976; Thornberry and Jacoby, 1979). Monahan (1984) reviewed these first generation studies and concluded that even the very best risk assessment technology was only able to get it right once out of every three attempts. He found the best predictors of violence among the mentally disordered were the same demographic factors that forecast violence among those who did not suffer from mental illness. He was particularly scathing about psychological indicators such as diagnosis or personality traits, which he claimed were the poorest predictors. Subsequent studies have challenged these conclusions. Taylor (1982) and Binder and McNeil (1988) have demonstrated links between rates of violent offending and specific clinical diagnoses. The MacArthur Violence Risk Assessment Study (VRAS; Monahan et al 2000) also highlights the significance of clinical factors such as substance abuse and psychopathy in the prediction of violent outcome among non-forensic psychiatric patients discharged from hospital. The current view is that we can enhance predictive efficiency by considering both static (historical) and dynamic (mental state, social circumstances, stressors, etc) factors (Monahan et al 2000), although high levels of accuracy are not necessarily crucial to the legal standing of risk assessment.

1.3 Violence prediction and the law

In 1974, the editors of the Harvard Law Review could write: "The difficulty of predicting the type and severity of an individual's anticipated antisocial conduct creates grave doubts as to the constitutional validity and the wisdom of present statutes authorising police power commitments"; (Developments in the Law, 1974, page 1245).

This concern about whether or not clinicians were so inaccurate at predicting violence that giving legal standing to their opinions had violated the 8th and 14th amendments of the constitution of the United States pervaded American legal discourse on the topic in the 1970s (Brooks, 1974). During the decade that followed, it became clear from decisions made in the US Supreme Court that even though the prediction of violence was imprecise, it still had a part to play in protecting public safety, and did not violate the founding tenets of the constitution.

A land mark case in this area was that of Thomas Barefoot (reported in Monahan, 1996), who was convicted of the murder of a police officer. At a separate sentencing hearing the jury considered the two questions put to it under the Texas death penalty statute; was the conduct causing the death committed deliberately and with reasonable expectation that the death of another would follow, and is there a chance that the defendant will commit further criminal acts of violence that would constitute a continuing threat to society? The jury answered, "Yes" on both counts, which required the imposition of the death penalty.

In Barefoot v Estelle (1983) the US Supreme Court considered the constitutionality of using clinical predictions of violence for the purpose of determining whom to execute. Justice White wrote "... it is urged that psychiatrists, individually and as a group, are incompetent to predict with an acceptable degree of reliability that a particular criminal will commit other crimes in the future and so represent a danger to the community. The suggestion that no psychiatrist's testimony may be presented with respect to a defendant's future dangerousness is somewhat like asking us to dis-invent the wheel.... If the likelihood of a defendant committing further crimes is a constitutionally acceptable criterion for imposing the death penalty, which it is, ... and if it is not impossible for even a lay person sensibly to arrive at that conclusion, it makes little sense, if any, to claim psychiatrists, out of the entire universe of persons who might have an opinion on the issue, would know so little about the subject that they should not be permitted to testify", (pp 896 – 897).

In Schall v Martin (1984) Justice Rehnquist stated: "... from a legal point of view there is nothing inherently unattainable about a prediction of future criminal conduct. Such a judgement forms an important element in many decisions, and we have specifically rejected the contention ... that it is impossible to predict future behaviour and that the question is so vague as to be meaningless.", (page 2417).

Not only courts, but professional organisations as well have concluded that predictions of violence are here to stay. The American Psychiatric Association's (1983) guidelines included the involuntary hospitalisation of people with mental

disorders who were seen as likely to cause harm to others. In the United Kingdom there is a clear parallel with the powers invoked under the Mental Health Act 1983.

Guidelines for involuntary civil commitment of the National Center For State Courts (1986) suggested that; "Particularly close attention be paid to predictions of future behaviour, especially predictions of violence and assessments of dangerousness. Such predictions have been the bane of clinicians who admit limited competence to offer estimates of the future, yet are mandated legally to do so. However, such predictions will continue to provide a basis for involuntary civil commitment, even amid controversy about the scientific and technological shortcomings and the ethical dilemmas that surround them", (page 493).

Where constitutional law once framed the legal questions asked of violence prediction, tort law (a breach of duty for which the person is held responsible) now takes precedence (Monahan, 1996). The landmark case in this area is Tarasoff v Regents of the University of California (1976). The California Supreme Court's ruling was that psychotherapists who know or should know of their patient's likelihood of inflicting injury on identifiable third parties have an obligation to take reasonable steps to protect the potential victim. Most courts, both in America and the United Kingdom, have accepted the essence of a duty to protect, and this has become a fact of professional life for nearly all clinicians. Attempts by practitioners to do their tasks well and avoid legal entanglements have led not only to the quest for reliable risk predictors, but also to further investigation of how violence comes about.

1.4 Theories of violence

Research on violence has been guided by theories which differ in their emphasis on learned or innate components, affective or cognitive processes, and internal or external determinants. The major competing assumptions are represented by biological, psychodynamic, social learning and social psychology perspectives (Blackburn, 1995).

Biological approaches propose that violence is dependent on inborn structural characteristics of brain and musculature. Their co-ordination is controlled by specific neurochemical systems, and show similarities between humans and other animals. Lorenz (1966) proposed a universal instinct of aggression, the purpose of which is population control, social organisation, selection of the strongest for reproduction, and brood defence. Instinct relates to a spontaneously generated energy source in the nervous system which discharges through fixed action patterns in response to specific releasing stimuli. The theory postulates that human behaviour is governed by a constant need to release aggressive energy.

Moyer (1981), drawing on studies examining the effects of electrical brain stimulation and surgical lesions in animals and in patients with organic pathology, suggested there are organised neural circuits in the brain. These are sensitised by hormones and blood constituents, and when fired in the presence of a relevant target produce integrated attack behaviour. Human learning can influence the selection of targets and the inhibition of behaviour, but feelings of hostility will be experienced whether or not aggressive action occurs. Along similar lines, Mark and Ervin (1970) proposed

centres in the limbic lobes of the brain which control aggressive behaviour, and that these are damaged in a substantial proportion of violent people.

Lorenz (1966) found support for his theory of instinctive aggressive drive amongst the ranks of psycho-analysts. They assume such a motivation and are interested in how it is accommodated within the hypothesised psychic structures of id, ego and superego (Blackburn, 1995). Manifestations of aggressive instinct include the rage reaction to frustration, which mobilises the organism for combat. Aggression may be provoked by external events, but instinctive aggressive impulses are believed to be constantly generated. They may erupt as irrational violence in pathological characters lacking superego control. However, aggressive energy may also be displaced to other targets, or purged through substitute expression (Catharsis).

The concept of an aggressive instinct in humans has not generally been accepted by behavioural scientists (Blackburn, 1995). Research has failed to uncover evidence in favour of the hydraulic model of a reservoir of energy, even in lower animals, or to support the idea of human brain systems exclusively concerned with aggression (Valenstein, 1976). It is true that organic damage to the brain can sometimes produce irritability and aggressiveness, but this appears to reflect non-specific disorganisation rather than interference with particular locations in the brain (Blackburn, 1995).

The same criticisms apply to the psychodynamic instinct model. The idea that many constructive activities are manifestations of somehow transformed destructive

energy allows virtually any action to be construed as aggression, and is not amenable to scientific testing (Blackburn, 1995).

While accepting the possibility of primitive connections between threatening events and motor responses, learning theorists reject the concept of an aggressive instinct (Blackburn, 1995). They hold that aggression is acquired and maintained by reinforcing contingencies. A violent response which successfully attains a desired reward is likely to be strengthened and repeated (positive reinforcement), as is one which removes a threat (negative reinforcement). Aggression followed by immediate punishment is less likely to be repeated. Cognitive mediators are important. For example, frustration and punishment are less likely to provoke aggression if they are perceived as justified.

The learning theory view that aggression originates primarily in family modelling and reinforcement is currently the most influential. Longitudinal studies show that family histories of violent children are characterised by higher rates of parental deviance or indifference, marital conflict, and lack of supervision than the norm, while violent adults frequently report witnessing aggression and experiencing abuse as children (Blackburn, 1995). The findings of the meta-analysis carried out by Bonta, Hanson and Law (1998) are consistent with social psychological perspectives of criminal and violent behaviour (Akers, 1985; Andrews and Bonta, 1994). These can be understood in much the same manner as any other learned pattern. Early childhood environment and social position provide the contexts in which criminal activities are modelled and reinforced (Bonta et al, 1998).

Bandura (1983) proposed a social learning theory of aggression. He saw it as originating in modelling and reinforcement, through which people develop expectations about the likely outcomes of different behaviours in achieving a variety of goals. Actions are adjusted to meet personal and social standards through the regulatory processes of reward and punishment. Criteria may be overridden or neutralised by cognitive distortions such as blaming or dehumanising the victim.

Bandura rejected the notion of a specific aggressive drive, and believed that both aversive experiences and positive incentives produce a general increase in emotional arousal. This motivates whatever relevant responses are strongest in the behavioural repertoire. In dealing with challenge, aggression is only one of several possible coping strategies. Others might be avoidance or constructive problem solving, depending on the individual's skills. Anti-social aggression may actually reflect a failure to learn more adaptive ways to resolve conflict.

According to Andrews and Bonta (1994), the social and personal factors that reward and fail to punish criminal activity are an established history of benefiting from it, a social environment that encourages and tolerates criminals, personal attitudes and values which support breaking the law, and a personality style that finds impulsive high risk behaviour rewarding. Further evidence for the social psychological view comes from the work of Gendreau et al (1996). They found that family problems, poor living arrangements and substance abuse were all related to recidivism.

Some social psychologists argue that aggression has to be understood in terms of the social context and meaning of a violent act to the perpetrator. Toch (1969) saw

threats to masculine self-image and the promotion of a particular view of the self as significant in violent encounters between police and delinquents. Similarly, Felson (1978) proposed that aggression was a means of restoring a person's threatened identity. Wolfgang and Feracutti (1967) wrote about a sub-culture of violence in urban ghettos. They identified a "machismo" pattern of attitudes favouring honour, status and masculinity which legitimised violence when these were threatened.

However, social and cognitive theorists propose that it is not just the perception of unpleasantness which instigates aggression, but rather the appraisal of an event as threatening and unjustified. Some examples are the impact of perceived iniquity on group violence and an attributional bias demonstrated by aggressive boys to assume the worst of their peers (Blackburn, 1995). What may appear to an observer as "unprovoked" violence might be the result of the aggressor's cognitive distortions.

Behavioural analysis of the chain of events leading up to a violent act suggests that some assaults may be precipitated by the conduct of the victim (Wolfgang, 1958). This is not an attempt to blame the victim for the outcome, but it is consistent with the view that aggression involves impression management, and study of homicide reports shows that retaliation is a significant factor (Wolfgang, 1958). This may be self-defence or an attempt to save face. It might be that a violent outcome is more the result of an escalating sequence of events than the characteristics or goals of the participants (Blackburn, 1995).

Environmental factors also have some part to play in the instigation of aggression. Studies show that hotter regions and warmer time periods are associated with

increased rates of murder, rape, assault, riot, and domestic violence (Anderson, 1989). The acts may be partly mediated by changes in social contact and opportunities for aggression, but the evidence is consistent with effects at the individual level (i.e. variability in mood, or misinterpreting the reasons for arousal).

Current theories emphasise the role of cognitive factors in the inhibition and facilitation of aggression (Blackburn, 1995). Intellectual deficits may play a part, along with poor social skills. In conflict situations, negotiation tactics must be used if violence is to be avoided. Slaby and Guerra (1988) found that violent delinquents and aggressive schoolboys showed more problem solving deficits than did non-aggressive male students. The first two groups also reported more positive beliefs and protective rationalisations about the consequences of aggression, claiming that violence made them feel better about themselves and tending to minimise the amount of distress their victims experienced.

The stability of aggressiveness as a personality characteristic has been established (Blackburn, 1995). A high correlation between initial assessments of aggressiveness and those obtained between six months and 21 years later was recorded by Olweus (1979) in his review of studies. Farrington (1989) found that almost a quarter of boys rated as highly aggressive by teachers at ages 12 – 14 subsequently had a conviction for violence. This is compared to a figure of 7% for less aggressive boys.

However, learning to resolve conflict through violence under the conditions which exist in a dysfunctional family is not an adequate explanation for the continuation of an aggressive style of interaction throughout life. Learning theory predicts that non-

reinforced behaviour will be extinguished. Continuity is accounted for in terms of expectancy confirmation processes (Blackburn, 1995). For example, a hostile person expects others to be hostile, and behaves in ways which elicit the expected reaction. This outcome confirms the cognitive distortions held about others, and strengthens the cycle. Opportunity for changing expectancies and learning alternative ways of coping are therefore minimised.

No one of the biological, psychodynamic, social learning, or social psychology approaches has all the answers to the riddle of why a particular individual behaves violently in a certain set of circumstances. As Blumenthal and Lavender (2000) conclude, manifestations of aggression come about through the complex interaction of many variables. A clearer understanding of the origins of violence might be possible if investigators spent more time refining research problems and constructing conceptual frameworks through which to view the questions they want their data to answer (Jackson, 1997).

1.5 The conceptual contours of violence

The specification of the basic facets of a problem and the clarification of the elements of those parts have been formalised in Guttman's technique of facet analysis (Guttman, 1957, 1971; Levy, 1985). Guttman developed this approach to deal with difficulties of definition that arise when constructing item sets to measure psychological concepts. The emphasis is on placing restrictions on the domain of enquiry by defining the phenomena of interest and their logical relationships. Although facet analysis is not well known to psychologists, it does provide a useful framework for posing precise research questions and making predictions.

The inquiry is stated in the form of a mapping sentence. In the case of predictions of violence, a mapping sentence serves the purpose of specifying the logical relationships between, and clarifying the content of, the basic units of the problem; "What factors provide information about who is going to do it again?" The mapping sentence (MS 1) set out in Figure 2 is an initial attempt to apply facet analysis to the prediction of violence (Jackson, 1997);

Figure 2 **MAPPING SENTENCE FOR QUESTIONS ABOUT THE PREDICTION OF VIOLENCE**

	Facet 1		Facet 2
	clinical judgement		are thought to index
	theoretical		predict
What	aftercare	factors	are related to
	situational		explain
	historical		cause
	demographic		

	Facet 3		Facet 4
	Individual		be accused of
Which/an	type of person	will/to	be convicted of
	Member of a group		commit
			be cautioned for
			be charged with

	Facet 5		Facet 6
	Parole violation		the parole period
	Criminal act		no specified period
a	Violent act	within	a period following parole
	dangerous act		time period X
	aggressive act		
	impulsive act		

MS1 consists of six conceptually independent facets. The first is a risk factor, the second specifies the extent of determination, the third distinguishes the population of interest, the fourth outlines the legal status of the act, the fifth identifies the type of act, and the sixth delimits the time of offence (Jackson, 1997). The elements of each facet restrict the domain of enquiry by defining the phenomena of interest, and allow the investigator to delineate the research problem by combining elements from each into a single sentence. For example, the principal question posed by my research dissertation might be framed as; "What mostly historical factors predict which individual will be cautioned for, charged with, or convicted of a violent act within time period x (i.e. two years of completing a KFPS risk assessment interview)". Knowing that aggressive behaviour is a complex phenomenon made up of many diverse but interactive components raises questions about whether or not there are valid ways to predict its occurrence.

1.6 The validity of violence prediction

It is not an easy task to summarise the work done on the accuracy of violence prediction (Douglas, Cox and Webster, 1999). The few studies that existed in the 1960's and 1970's have been referred to in the section of this portfolio headed, "A critical review of the psychological literature on the controversy between clinical and actuarial approaches to decision making". There was "clear and convincing evidence" that mental health professionals were unable to predict violence accurately (Cocozza and Steadman, 1976, page 1084). Only one in three positive predictions of violence were accurate (Monahan, 1981). However, more recent reviews have concluded that due to better methodology and techniques of statistical analysis, accuracy has improved.

Studies have shown consistently that certain diagnoses, symptoms and previous aggressive behaviours are related to violence among psychiatric inpatients (McNiel and Binder, 1991, 1995; McNiel et al, 1988).

Several domains of variables (i.e. quality of early family life, current intimate relationships, arrest history, admissions history, and assault as part of the presenting problem) were found to predict violence by Klassen and O'Connor (1989). Using these indicators the authors were able to classify correctly 76% of patients into violent or non-violent groups.

The McArthur risk assessment project has led to the development of a violence prediction tool that relies upon a regression tree approach. The clinician asks certain questions in a predetermined order, and considers risk markers depending on the

answers the patient has given. This approach has successfully enabled the authors to divide psychiatric patients into high or low risk categories (Monahan et al, 2000; Steadman et al, 2000).

Using four variables (i.e. young age, three or more violent acts, heavy drug use, and hostility) Gardner et al (1996) developed a screening instrument for use with psychiatric patients. With a regression tree design similar to that used by the McArthur group, the authors were able to show excellent specificity (i.e. the proportion of persons who are not violent and who were predicted to be not violent) at 99%. However, the sensitivity of the device was poor at only 7% (i.e. the proportion of persons who are violent and who were predicted to be so).

One promising and robust finding is the relationship between psychopathy, as measured by the Hare Psychopathy Checklist (revised and screening versions), and violence. Two meta-analyses have shown this link to be at least moderate (Hemphill, Hare and Wong, 1998) to large (Salekin, Rogers, and Sewell, 1996).

Mossman (1994) determined that predictive accuracy had improved with recent studies, that actuarial approaches to decision making about risk out- performed clinical ones, and that there were data sets available which gave some cause for optimism. However, the old question, "Can violence be predicted?" was no longer appropriate. It might be more productive to adopt the facet analysis approach put forward by Guttman (1971) and refined by Jackson (1997), and ask, "Which groups of subjects, with which particular characteristics, followed over what periods of time,

are likely to exhibit precisely defined kinds of violent behaviour?" (Douglas et al, 1999; page 161)

In order to answer such a refined research proposal it is necessary to consider the impact of statistical characteristics of the populations under study, and ways in which the data produced might be dealt with.

1.7 Statistical measures for assessing accuracy of prediction

No discussion of violence prediction and its accuracy can afford to leave out the issue of base rates (i.e. the normal frequency of occurrence of violence in a particular population or group; Blumenthal and Lavender, 2000, page 41). This data is used by researchers to evaluate the effects of specific manipulations(Reber, 1985).

A person in a given area, or from a particular group in a population, will be more or less likely to engage in acts of violence than someone in another place with different base rates of violence, whether or not they exhibit known risk markers (Blumenthal and Lavender, 2000). Regional variations in violence in the USA have been commented on by Monahan (1993). In general, smaller communities have a lower prevalence of violence, and within the same city, neighbourhoods can differ dramatically.

Ignorance of the relevant base rate for a behaviour can give an extremely misleading picture (Moore, 1996). For example, if a hypothetical news item stated that escapes from Kent Prisons had doubled over the last 10 years, the reader may well be outraged. This view might be modified if the article went on to explain that in the current decade there had been just two escapes, and only one in the decade before that.

Sometimes distortion can arise from quoting statistics from an unhelpful angle. A statement on urban crime suggested that the majority of muggings in London were perpetrated by young Afro-Caribbean men (Moore, 1996). Whether or not this is

true, a more helpful statistic would be the proportion of young Afro-Caribbean men who offend in this way (a minority of that sub-population).

A general conclusion reached by Harris, Rice and Quinsey (1993) is that, "The data on the prediction of violence indicate that clinicians are insensitive to variations in the base rate of violent offending, show poor agreement among themselves, make the same judgements as lay persons, and are less accurate than actuarial models", (p 332).

This depressing state of affairs has led researchers to look for more accurate, empirically based ways to forecast violence. Instruments such as the Violence Risk Appraisal Guide (VRAG), the Psychopathy Checklist – Revised (PCL-R) and the Historical/Clinical/Risk Management 20 - item scale (HCR-20) are gaining recognition as reliable and valid means of estimating risk (Dolan and Doyle, 2000). Along with the development of these scales has come awareness that statistical techniques other than correlation and regression are needed to compare accuracy (Douglas et al, 1999). The advantages of receiver operating characteristic analyses (ROC) over other techniques have been outlined by Mossman (1994). They are much less dependent on base rates of violence, and produce a statistical index called the area under the curve (AUC). This is derived from plotting sensitivity (true positive predictions on the vertical axis) against specificity (true negative forecasts on the horizontal axis) for various cut-off scores on a predictive measure. The AUC can range from zero or perfect negative prediction to 0.5 which is no better than chance, on to 1.0 or perfect positive forecasting. Values for AUC in the mid 0.70s and upwards may be considered large effects, and represent the probability that a

randomly chosen, actually violent person will score higher on a particular risk prediction instrument than a randomly chosen, actually non-violent person (Douglas et al, 1999). It is anticipated that the reporting of ROC data will make comparisons between studies of the accuracy of violence prediction instruments more empirically sound (Dolan and Doyle, 2000).

A sobering finding by Steadman (1983) is that the likelihood of violent recidivism among a group of offenders is so low that it is difficult to improve on the accuracy of the forecast that nobody will re-offend, even in high risk groups. However, the cost of such actions is so great that mental health professionals are obliged to make an effort to identify individuals who are at risk of committing violence, so that proper steps can be taken to protect the community. Originally, this was based on unaided clinical judgement, but a growing research literature has found fault with this, and lead on to greater refinement in violence risk prediction.

1.8 Approaches to violence risk prediction

Historically, those attempting to forecast violent behaviour have based their decisions on straightforward clinical judgement. Flaws in this process have lead to the development of actuarial methods, and a division amongst practitioners about how best to proceed. This argument has been examined in another part of my doctorate portfolio. A compromise solution has come about in the form of structured clinical judgement.

Each of these approaches will now be considered in more detail.

(a) Unaided clinical risk assessment.

In clinical practice, assessments of the risk of danger or violence an individual poses are usually based on unaided clinical judgement. There are a number of disadvantages to this approach. Some are low inter-rater reliability, low validity, a failure to clarify all the steps in the decision making process (Monahan and Steadman, 1994; Webster et al, 1997), and inferior predictive validity compared to actuarial forecasts (Meehl, 1954: Lidz et al, 1993; Mossman, 1994).

However, the clinical method is not without some merit. Gardner et al (1996) showed that while actuarial measures were better than clinical ratings at identifying repetitively violent patients who suffered from mental illness, the clinical ratings were still better than chance.

Predictive accuracy can be enhanced when clinicians consider the context in which violence occurs among their patients (Mulvey and Lidz, 1985). Consensus among multidisciplinary team members about predictions of risk has been found comparable with actuarially based schedules over similar time-scales (Fuller and Cowan, 1999).

(b) Actuarial methods

Using this approach, assessors arrive at decisions based on data which can be coded in a clearly defined manner (Meehl, 1954). Judgements are made according to rules, and focus on risk factors that have been empirically proven to predict violence in particular settings and with specified populations. Some of these markers are static (e.g. demographic variables) and can be identified from historical records by people without extensive clinical training or experience.

Although this method undoubtedly improves the consistency of risk assessment (Dolan and Doyle, 2000), some have argued that it ignores individual variations in risk, over emphasises static variables, fails to prioritise clinically relevant markers, and minimises the role of professional judgement (Hart, 1998; a, b). These criticisms are challenged in another part of my doctorate portfolio.

Despite such concerns, actuarial risk assessment tools have been accepted for some time in American prison settings when parole decisions fall due. Examples include the Base Expectancy Score (Gottfredson and Bonds, 1961), the Level of Supervision Inventory (Andrews, 1982), the Salient Factor Score – Revised (Hoffman, 1983), and the Statistical Information on Recidivism (SIR) scale (Nuffield,

1989). In the UK, similar measures have been developed to give risk of reconviction scores for prisoners being considered for parole (Copas et al, 1996).

(c) Structured clinical judgement

This is a compromise position between unaided clinical expertise and strict adherence to actuarial procedures. The leading proponents of this model (Webster et al, 1997b) argue that risk prediction can be substantially improved if:

(i) Assessments are carried out using clearly defined protocols which have been critically reviewed in the research literature.

(ii) Agreement between assessors is good, with commonly shared standards of training, knowledge, and expertise.

(iii) Prediction is for a defined type of behaviour and over a specified period.

(iv) Violent acts are detectable and recorded.

(v) The information necessary to make informed decisions is available to the assessors and has been substantiated.

(vi) Actuarial estimates are adjusted only if there is sufficient justification.

Harris et al (1993) view the above steps as "Structuring discretion", where clinical judgement can be improved through the integration of actuarial information. Clinical decisions are anchored by actuarial estimates of risk.

Psychometric instruments have been developed with these considerations in mind for assessing risk of violence in clinical contexts. They include the

Historical/Clinical/Risk Management 20 – item scale (HCR-20; Webster et al, 1997b), the Spousal Assault Risk Assessment Guide (Kropp et al, 1995), the Sexual Violence Risk (SVR – 20) scale (Boer et al, 1998), the Violence Risk Appraisal Guide (VRAG; Harris et al, 1993), and the Psychopathy Checklist – Revised (PCL- R; Hare, 1991).

Some of these merit closer examination.

1.9 Violence prediction instruments used in clinical settings

(a) The Violence Risk Appraisal Guide.

The VRAG (Harris et al, 1993) is made up of 12 items scored on the basis of a weighting procedure developed on a calibration sample of 618 males charged with severe violent crimes. Using ROC analysis, Rice and Harris (1995) found the VRAG predicted violent recidivism with AUC's of 0.75, 0.74, and 0.74 for 3.5, 6, and 10 years respectively.

Reports suggest this instrument is less valuable in predicting violent sexual recidivism in paedophile sex offender populations (Rice and Harris, 1997). It has been criticised because of its reliance on relatively static factors, and Webster et al (1994) recommend that it be supplemented with a checklist of clinical features to produce a violence prediction scheme.

(b) The Historical/Clinical/Risk Management 20 – item scale

The HCR-20 (Webster et al, 1997b) contains ten historical items (two of which address the issue of personality dysfunction), five clinical inquiries, and five risk management markers. It shows good inter-rater reliability (Webster et al, 1997a). While studies are limited to the work of a small group of North American researchers, the data generally show better than chance relationships between HCR – 20 scores and violent outcomes (Dolan and Doyle, 2000).

Douglas, Cox, and Webster (1999) found that across civil psychiatric, mentally disordered offender, and correctional service samples, the HCR – 20 showed effect sizes with violence that typically are either large or moderate. They report that in the few studies that have compared its validity to other measures (e.g. VRAG, PCL-R, PCL:SV) it has performed well. Reliability has ranged from acceptable to good, with most inter-rater indices in the 0.8s. Internal consistency has been documented in the high 0.70s to the 0.90s (Douglas, Cox, and Webster, 1999).

(c) Psychopathy Checklist (Revised).

The 20 item PCL-R (Hare, 1991) was originally devised as a research tool for operationally defining the concept of psychopathy. Scores range from 0 to 40, and those who produce a total of more than 30 are seen as demonstrating characteristics consistent with a diagnosis of psychopathy.

The PCL-R has been shown to have good psychometric properties (Cooke, 1998), and a number of studies have demonstrated its utility as a risk assessment tool in identifying recidivists, and predicting violence in North American forensic and prison populations (Hart, 1998a).

Grann et al (1999) found that for Swedish offenders with personality disorder, PCL-R scores were the best predictor of violent recidivism two years after release.

The PCL-R has been shown to be superior to other classical actuarial risk scales on indices of general recidivism and violent re-offending (Harris et al, 1993; Rice and Harris, 1995; Zamble and Palmer, 1996; Hemphill et al, 1998).

Hemphill and Hare (1996) compared the predictive validities of the PCL/PCL-R and several actuarial measures. The finding was that they performed similarly for general recidivism, but the PCL-R was significantly better at forecasting violent re-offending.

The PCL-R has been supplemented by a 12 item screening version called the PCL-SV (Hart et al, 1995). It has similar psychometric properties to the PCL-R (Dolan and Doyle, 2000). Scores range from 0 – 24, with those producing totals above 18 satisfying the requirements for a diagnosis of psychopathy.

Currently the PCL-R and PCL-SV are believed to be some of the most reliable measures for assessing personality constructs likely to be relevant to violence prediction (Dolan and Doyle, 2000).

Despite the now generally accepted utility of all the above measures when it comes to predicting recidivism and violent offending, there is still widespread reluctance to use them. Some of these qualms are discussed elsewhere in this portfolio in the section dealing with a critical review of the psychological literature and the controversy between clinical and actuarial approaches to decision making. There is, however, one factor often overlooked when conducting a risk assessment, and that is the amount of practitioner's time involved in carrying it out. The psychometric

instruments mentioned involve checking through usually extensive files about the patient, conducting an interview, often seeking out collateral information from others who know the person, and synthesising the whole into a report. It is not always easy to do this promptly and properly, which has led the writer to the doctoral research project. Perhaps fast and frugal reasoning using probabilistic mental models (PMM; Gigerenzer and Goldstein, 1996) can produce a risk assessment heuristic which matches the PCL-R and PCL-SV in terms of reliability and accuracy, and yet involves fewer resources. Investigation of the applicability of PMM's has not, as yet, turned to the prediction of violence, but when it does the research carried out will need to satisfy certain criteria.

1.10 Characteristics of good research and routes for future inquiry

If investigation of violence prediction is to advance the state of science, it must reflect the following characteristics (Monahan and Steadman, 1994):

- Dangerousness must be broken down into its component parts; the variables used to predict violence (risk factors), the amount and type of violence being predicted (harm) and the likelihood that harm will occur (risk).
- A variety of theoretically based risk factors should be chosen.
- Harm needs to be scaled in terms of seriousness and assessed through many different measures.
- Risk must be treated as a probability estimate that changes with time and context.
- Priority should be given to actuarial research that attempts to establish a relationship between risk factors and harm.
- Patient samples should be large, representative, and from many different settings.
- Managing as well as assessing risk must be the goal.

After reflecting on this list, one of the authors decided to add risk communication to it (Monahan, 1996). He sees this as an essential adjunct to risk assessment.

The National Research Council (1989) defines risk communication as, "An interactive process of exchange of information and opinion among individuals, groups and institutions; often involving multiple messages about the nature of risk or

expressing concerns, opinions, or reactions to risk messages or to legal and institutional arrangements for risk management", (page 322).

In the past, predictions were often given in Yes/No terms, where an individual was deemed "dangerous" or "not dangerous" (Douglas et al, 1999). Such an approach might be seen as too simple and increases the likelihood of error. Some researchers advocate communicating violence risk estimates in terms of precise probability levels (Quinsey et al, 1998), while others prefer more general statements such as low, moderate, or high risk (Monahan and Steadman, 1996; Webster et al, 1997b). There is support for the latter view, with some arguing that the specification of precise levels of probability conveys a misleading sense of exactitude (Douglas et al, 1999).

Risk of violence is conditional upon a great many factors that may occur after an assessment and prediction have been completed (Mulvey and Lidz, 1995). Evaluation of such contingencies is not a static process, and the degree of dynamism concerned precludes the possibility of giving fixed estimates (Douglas et al, 1999). The authors illustrate this with the following example:

"If person A returns to substance abuse (a known risk factor) while person B receives appropriate treatment by diligent and trained professionals and is reunited with a supportive and strong family, the probabilities may not remain the same. When this idea is coupled with another concept – that of relative risk – the form of violence risk communication that is considered appropriate includes categories of low, moderate, high and imminent (see Monahan and Steadman,1996), that are conditional upon (may be aggravated or mitigated by) X and Y factors and are

relative to other inmates/patients/insanity acquittees with similar demographic characteristics from similar institutions", (pages 175 and 176).

The concept of risk management integrates some of these ideas. For example, on the basis of information known about a person they may be considered to pose a moderate risk of violence when seen in relation to other patients released from the same institution. The assessor's judgement is clearly conditional upon risk factors staying the same. However, if the person stops taking their medication, absconds from hospital, or returns to using illicit substances then the opinion may shift towards high risk (Douglas et al, 1999). The advantages of such an approach are the implications for management that follow the assessment (Boer et al, 1998), and the flexibility to change communications about violence risk if conditions alter.

Monahan (1996) concludes that an important task for those working in the area of violence prediction is to find ways of communicating the results of risk assessments to relevant audiences so as to facilitate risk reduction and management.

Webster et al (1995) state that published findings have stressed the importance of simple actuarial variables over complex clinical ones. They believe that future research should break the task of assessment into three general areas. The first has to do with historical or background factors. The second involves face-to-face clinical evaluation, paying attention to variables such as insight, attitude, symptoms, stability, and likelihood of responding to treatment, but without overweighting them. The third involves estimation of risk factors such as the feasibility of the patient's discharge plans, access to victims, social support, compliance with medication and other

treatments, and reaction to environmental stressors (Webster et al, 1997b).Statistically significant relationships exist between historical, clinical and risk factors and threatened violence (Webster et al, 1995). The task ahead of scientists and practitioners is to see if these correlations can be improved upon, and how they might form a part of the decision making process about individual cases.

Douglas et al (1999) maintain that researchers should help to consolidate the scientific literature in ways that have practical use to clinicians, and go some way to bridging the divide between actuarial and clinically based decision making. There seems much to be gained from a plan of action that draws clinicians and researchers together in the common interests of making decisions that affect so profoundly the lives and freedoms of people suffering from mental and personality disorders who are in trouble with the law.

In the United Kingdom, there have been few high quality studies, and conclusions have been limited by the choice of measures of violence and lack of adequate control groups. Blumenthal and Lavender (2000) believe that a large scale epidemiological survey in the UK using multiple measures of violence could make an important contribution to the literature.

Link and Stueve (1995) recommend the epidemiological cohort design which specifies the mental disorders of interest, uses representative samples of individuals with no history of mental disorder as community controls, and compares the groups' subsequent involvement in violence.

Substance abuse is one of the most robust predictors of violence. Future research might include investigations of which symptom patterns combine with this to produce aggression (Monahan, 2000).

The postal survey of consultant psychiatrists carried out as part of this doctoral dissertation revealed that non-compliance with treatment by psychotropic medication was number 6 in a list of cues looked for as part of attempts to predict future aggression. However, to date there has been no prospective research on any link between non-compliance with treatment and future violence (Blumenthal and Lavender, 2000).

Instead of studying categories of disorder (such as schizophrenia), it might be more productive to investigate temperaments, psychological traits such as impulsivity, and active symptomatology (Marzuk, 1996).

Further research is needed on the link between violent thoughts and subsequent assaults. There is also a requirement for a typology of violence which differentiates between unique impulsive acts and more enduring aggressive behaviour (Blumenthal and Lavender, 2000). Both of these issues test professionals making predictions to the limit.

1.11 Directions for clinical practice and policy

(a) Clinical practice

As mentioned earlier in this chapter, there is a need to change from decisions based on unstructured, clinical experience-related impressions towards a more objective, transparent approach. Beginning with risk assessment, practitioners should move on to risk management, and then to a consideration of wider service strategies (eg. operationally defining a satisfactory information base, agreeing on the form proper documentation should take, and identifying ways to respond to crisis). Monahan (1993, b) describes this process as "risk containment", and it has implications for the way in which mental health professionals do their work.

Gunn (1993) recommends a seven step procedure to assess risk, beginning with a detailed personal history, and moving on to evaluations of substance use or abuse, sexual interests and attitudes, criminal or anti-social behaviour, psychological assessment of intelligence and personality, mental state, response to treatment, and a description of behaviour and attitudes.

Blumenthal and Lavender (2000) would add two further areas to this approach. The first is an acknowledgement of the social context of the person under evaluation, and those they are most likely to interact with. The second is a psychological formulation providing an analysis (based on theory and research) of all the above information collected during assessment. The purpose is to allow a fuller understanding of the history and development of the individual's difficulties which takes account of their

social and cultural setting. In this way the meaning of the violent act might become clearer.

When a formulation is complete it becomes possible to identify which risk factors are changeable and which are not. One is then in a position to suggest ways the former might be altered through treatment and management so that the likelihood of an undesirable outcome is lowered (Blumenthal and Lavender, 2000).

Management of risk involves the provision of security, supervision and support to lessen the impact of difficulties identified during the assessment phase. It is defined as the process of decision making and planning related to specific individual patients (Snowden, 1997), and is made up of medical, psychological, and social strategies to reduce the frequency and severity of problem behaviour.

Although there is clearly a need for risk assessment and management, it is important to reiterate that homicide is a rare event. In a population of individuals identified as being high risk, few will commit violence, particularly homicide. Focusing mental health resources on the issue of risk may result in other pressing priorities being ignored. There is, therefore, a need to keep a sense of perspective when attempting to predict and manage violence. It does the individual and their community no favours to over-commit time and money to the imprecise art of forecasting conduct at the expense of providing more mundane (and less dramatic) services to a larger group of patients. Appleby et al (1999) reported that standards of clinical practice could be improved by paying more attention to training of mental health staff, communication of opinions, using newer types of psychotropic medication to

enhance compliance, recognition of the part alcohol and drug abuse play in facilitating violence, taking seriously the feedback from care givers about threatening or irritable behaviour as a prodromal symptom, and the development of audit measures with relation to risk management. None of these recommendations are prohibitively expensive in terms of time, resources or money, though without such changes there remains an unacceptably high likelihood that potentially violent patients may be lost to follow up, and go on to commit acts which blight their lives and the lives of those around them.

(b) Policy

The overall level of violence in society is barely affected by those suffering from mental disorder (Monahan, 1992). When Wessley (1993) used Epidemiological Catchment Area data to calculate the proportion of violent incidents attributable to mental illness, he found this was only 3% of the total. Shaw et al (1999) comment on the small number of individuals with a mental illness who committed homicide in the UK identified by the National Confidential Inquiry. Taylor and Gunn (1999) have shown there has been little change in the contribution of mentally ill people to the number of homicides in England and Wales over a 38 year period. Despite these sobering figures, the subject is debated in a way that other correlates of violence are not. Young men are not stigmatised, yet youth and gender are more potent predictors of violence than mental disorder (Blumenthal and Lavender, 2000). However, the incidents which do occur are well publicised, often producing distress amongst members of the public and sometimes leading to demands for policy change restricting basic freedom. This has implications for the three statutory

agencies in the UK which are likely to become involved when matters of risk and propensity for violence are identified; the National Health Service, Social Services, and the Criminal Justice System (Blumenthal and Lavender, 2000). Working together has been difficult for these groups over the last two decades because of the amount of organisational change they have been required to manage (Timmins, 1996). The most significant development within Mental Health Services has been a shift in emphasis from care based in large institutions to a more community orientated approach (Blumenthal and Lavender, 2000). Difficulties in managing some mentally disordered people who have committed violent acts have resulted in ministerial pronouncements that community care has failed (Burns and Priebe, 1998).

British government policy has reflected an awareness of the need for better community services (Department of Health, 1995; Department of Health, 1996). Community Mental Health Teams (CMHT) have been created to provide continuity of care between inpatient and outpatient treatment. There has been criticism about the failure of CMHTs to operate effectively (Onyett and Ford, 1996) for people with severe mental disorder, and there is no consensus about whether or not this forms a core part of their work (Blumenthal and Lavender, 2000). Such a purpose needs to be clarified in the operational policies of CMHTs and to be acknowledged by team members.

Crucial to the work of these teams is the development and maintenance of good relationships between service users and staff (Onyett, 1999). If this does not happen

teams and services lose their cohesion and become fragmented, to the benefit of nobody (Blumenthal and Lavender, 2000).

Among policies which aim to provide for the delivery of effective community care is the Care Programme Approach (CPA; Department of Health 1990). This is a system for ensuring seamless delivery of care across all levels of patient involvement with statutory agencies. Its purpose is to plan for the discharge of psychiatric inpatients by developing and reviewing community care plans. This involves the assessment of need (including risk), which leads to the identification of tasks which must be completed in order to reduce the chances of relapse. A key worker is nominated to carry out certain activities in the therapeutic package, but primarily to organise regular review and monitoring of treatment plans. Such overview has been lacking in some instances where people with severe mental health problems have committed homicide (Shephard, 1996; Reith, 1998). Members of CMHTs are the most logical choice for the role of key worker, and are in the best position to ensure continuity of care across inpatient and community settings, as well as liaising with other services such as day hospitals, general practitioners, education, housing, and benefit departments (Blumenthal and Lavender , 2000).

The effectiveness of CPA has not been investigated scientifically, and there are concerns amongst mental health practitioners that the increased administrative workload has not been matched by obvious benefits to service users or staff (Blumenthal and Lavender, 2000).

The National Confidential Inquiry's recommendations stress the need for mental health trusts to have written policies on non-compliance with and disengagement from treatment (Appleby et al, 1999). It was felt that local homicide inquiries should no longer be held because they foster a culture of blame. Geddes (1999) viewed inquiries as methodologically inadequate for arriving at recommendations about mental health provision. They are retrospective, and foster the simplistic notion that homicide and suicide are preventable.

A Health of the Nation Strategy specifically highlighted the needs of mentally disordered offenders in the early 1990s, encouraging the expansion of court diversion schemes so that people would be dealt with by the service agencies most appropriate for the task (Performance Management Directorate, 1993). Despite this initiative the number of mentally ill in prisons remains substantial (Blumenthal and Lavender,2000).

There are proposals being considered in the UK to allow the preventative detention of individuals who are considered dangerous but who do not necessarily have a conviction for violence. The purpose of such policy would obviously be protection of the public, and this raises significant civil liberty issues (Eastman, 1999). It has been argued by Geddes (1999) that since homicide (and violence) are rare events, any increase in restriction would affect many people who would not go on to behave in an anti-social fashion.

The report of the National Confidential Inquiry (Appleby et al, 1999) estimates that improving compliance by introducing Community Treatment Orders may prevent two homicides per year, whilst financial and humanitarian costs would be considerable.

Policies for the management of the problems posed by those suffering from personality disorder have been shaped by this kind of reasoning (Appleby et al, 1999). This client group is the largest diagnostic category involved in episodes of homicide, and is strongly associated with violence in general (Blumenthal and Lavender, 2000). Such people often have previous convictions for violence, and tend to drop out of treatment. The Mental Health Act (1983), its subsequent revisions, and the Care Programme Approach do not seem to meet the needs of the personality disordered, who regularly form the basis of dispute between health staff and those employed by the Criminal Justice System about who should be providing oversight. Care and containment in the community is inappropriate for those who pose serious risk of violence, while imprisonment on its own does little to change the underlying causes of anti-social behaviour. However, there has been a development in the UK prison system where inmates are encouraged to address the reasons for their offending. HMP Grendon at Grendon Underwood in Buckinghamshire is run as a therapeutic community (Blumenthal and Lavender, 2000). Staff and residents interact democratically, with the latter being given much more responsibility for the daily running of the therapeutic wings than is usual in prison. When jails run as therapeutic communities were compared to standard institutions, Dolan and Coid (1993) found fewer symptoms of psychological distress, and a smaller number of serious incidents among violent offenders. Over all the years it has been operating,

HMP Grendon has only had one hostage incident, one escape, and no riots (Wilson and Ashton, 1998). It is noteworthy that in contrast to other prisons there is a waiting list of around 200 people who want to transfer there. Empirical support for the impact HMP Grendon is having on those who complete treatment may be found in a study by Marshall (1997). He followed up a cohort of 700 inmates admitted between 1984 and 1987. They had a lower probability of re-conviction than those on a waiting list control group. Another finding was that the longer people stayed at HMP Grendon the lower their rates of re-conviction.

On a cautionary note, however, the long term outcome remains uncertain, and results are conflicting. Among psychopaths who underwent compulsory treatment in one therapeutic community, violent recidivism was found to increase (Rice, Harris and Cormier, 1992).

1.12 Finishing comments

Violence prediction will never be entirely accurate, given that violence itself is a complex concept (Dolan and Doyle, 2000). This view is supported by Monahan (1996), who points out that decisions sending people to prison or to hospital (involuntarily) are often based upon untested assumptions about violent behaviour and the ability of professionals to anticipate it. However, some aggressive offenders are changed by exposure to treatment, and not all become recidivists. The therapeutic nihilism of "nothing works" can be countered with the more realistic argument that we do not yet know what makes a positive difference with violent criminals (Blackburn, 1995),

When Bonta et al (1998) showed that the major predictors of recidivism were the same for mentally disordered offenders as for non-disordered offenders, the suggestion was that risk assessment could be enhanced by paying more attention to the social psychology and criminology literature, while placing less reliance on models of psychopathology. This position may be seen as threatening by clinicians, who despite the lack of clarity about violence prediction are nevertheless frequently placed in situations where they are called on to make forecasts. These decisions often assume the form of "Yes" or "No" judgements, because they involve taking or not taking a particular course of action. Examples are releasing or not discharging a patient, and warning a third party of imminent danger or declining to do so (Mossman, 1994). However, accuracy in discriminating violent from non-violent people is only one factor effecting decisions. The threshold at which a rational judge takes action is influenced by the base rate of violence in the population group under

consideration, and a trade-off between the risks and benefits of various outcomes (Mossman and Somoza, 1992). A statistical procedure has been developed which is unaffected by underlying base rates and decision biases favouring certain prediction outcomes. Investigators can use Receiver Operating Characteristic analysis to evaluate the accuracy of violence predictions. However, clinicians and the general public need to realise that a proportion of judgements will inevitably be mistaken (Mossman, 1984). Acting under real world constraints decision makers will have to choose what kind of failures they can be most comfortable with. Faced with the prospect of hospitalising someone involuntarily or not, the proportion of false-positive and false-negative errors they make will depend on how they feel about the consequences of erroneously detaining someone who is not violent, against the results of discharging someone who is. The only way out of this dilemma is to obtain information about decision strategies and predictive accuracy that clarifies rather than muddles clinical and legal decision making. Fast and frugal reasoning using models of bounded rationality may provide a solution (Gigerenzer and Goldstein, 1996).

Chapter 2

Decision Making Strategies

2.1 The classical view and its alternative

Over the years scholars have defended the belief that people make judgements based on the laws of probability and statistics (Gigerenzer and Goldstein, 1996). It has been contended that decisions are arrived at after all possible alternatives have been surveyed, probabilities and utilities associated with every outcome estimated, and the value of each calculated, resulting in the alternative that scores highest being chosen (Simon, 1956, 1982; Michon and Pakes, 1995). Psychological research has assumed that statistical techniques are models of how we arrive at conclusions (Gigerenzer and Goldstein, 1996). Bayes' theorem, concerned with the calculation of conditional probabilities, has been seen as illustrative of human reasoning and memory (Anderson, 1990; Van Koppen, 1995). This view has been challenged by the notion that inference is biased and prone to error (Kahneman, Slovic, and Tversky, 1982). In an earlier work, Kahneman and Tverski (1973), the authors maintained that: "In making predictions and judgement under uncertainty, people do not appear to follow the calculus of chance or the statistical theory of prediction. Instead, they rely on a limited number of heuristics which sometimes yield reasonable judgements and sometimes lead to severe and systematic errors", (page 237).

When dealing with the complexity of the real world the human mind does not act as some kind of super-calculator, endlessly tallying in rational sequence all the

information necessary to arrive at a result. It complies more with a fast and frugal reasoning approach, where time and energy are limited, and the search for a conclusion is often stopped when the first compelling piece of information is reached (Gigerenzer, 1993; Gigerenzer and Goldstein, 1996). The term "satisficing" has been coined for this process (Simon, 1982; 1995). It is used to describe heuristics (procedures for solving problems) that successfully deal with conditions of limited time, knowledge, or computational abilities. This introduces the concept of bounded rather than classical rationality (Simon, 1955,1982). Organisms do not live in an ideal world where they have limitless resources at their disposal when faced with choice, and must make do with what is reasonably available.

Bounded (as opposed to perfect) rationality has cognitive and ecological components (Gigerenzer and Goldstein, 1996). Minds are adapted to the environment around them.

This insight that the minds of living systems can be understood in relation to the environment from which they evolved has had little impact on research into human decision-making (Gigerenzer and Goldstein, 1996). Simple psychological heuristsics have often been dismissed without a fair trial because they looked so unsophisticated by the normative standards of classical rationality. This is surprising when one considers that scientists have known and used similar principles since 1894 (Wilkening, 1973). Morgan's canon combated anthropomorphic thinking about animal behaviour by maintaining that: "In no case is an animal activity to be interpreted in terms of higher psychological processes, if it can be fairly interpreted in terms of processes which stand lower in the scale of psychological evolution and

development," (page 108). This doctrine strongly influenced behaviourism and is said to have come from Occam's razor (Wilkening, 1973), which maintains that concepts, principles, or factors should not be multiplied unnecessarily. The original Latin is "entia non sunt multiplicanda praeter necessitatem", (page 117). It is attributed to William of Occam (1280? – 1349?), an English political theorist and philosopher. He supposedly expressed the view that assumptions made about anything should be reduced or cut to a minimum (as if by a razor) because of his resentment at the medieval scientists' tendency to explain simple things ad infinitum. There are parallels with the law of parsimony, where, "In considering scientific formulations and constructs, the simplest approach is the best, all other things being equal", (Wilkening, 1973, page 97). Similarly, the law of least effort asserts that an organism will expend only a minimal amount of energy in exploratory behaviour in order to survive. These tenets are consistent with views put forward by Simon, Gigerenzer and Goldstein in their expositions of fast and frugal reasoning and bounded rationality. Awareness of them has contributed to the development of probabilistic mental model theory (Gigerenzer et al, 1991).

2.2 Probabilistic mental models

The theory of probabilistic mental models (PMMs) assumes that inferences about unknown aspects of the world are based on probability cues (Brunswik, 1955). The organism's decision strategy is influenced by the limited trustworthiness of signs it observes in the environment around it. This forces a probabilistic rather than a certainty seeking approach to judgement, where increased accuracy and efficiency are possible through accumulating knowledge in memory, and combining cues. Brunswik (1955) illustrated this using what he called the lens model, which represents the basic unit of psychological functioning, where cues are utilised according to their relevance and cost. However, the decision making process for the individual is not a perfectly transparent or clearly defined one. The chain from distant variables to nearby cues and on to central responses is made up of partial (or probable) causal explanations. Organisms are forced to behave as if in a semi-erratic ecology. This may be explained by drawing a comparison with signal detection theory in the field of telecommunications (Shannon and Weaver, 1949). Perceptual cues and behavioural means are like signals in coded messages (Brunswik, 1955). Mediating channels are contaminated with interference such as background noise, static or clutter, so that it is impossible to reconstruct a sent message with certainty no matter what manipulations are carried out. Whenever the capacity of a channel is less than the richness of variability of the source from which it accepts messages, the channel becomes overloaded. No matter how diligently engineers apply themselves to this problem there always remains some uncertainty about what the message was. The result is equivocation. Decisions must be made and action taken in the midst of confusion. When there are missing cues, the

cognitive system must make do by substituting whatever pieces of information happen to be available. Brunswik (1955) called this process vicarious functioning.

Gigerenzer and Goldstein (1996) propose a different view of Brunswik's lens model: "In a one-reason decision making lens, the first discriminating cue that passes through inhibits any other rays passing through and determines judgement," (page 665).

Some implications from this are that inference must be considered with regard to the real world context it occurs in, and that most probably organisms use satisficng heuristics along with knowledge of event frequencies in a variety of reference classes to make decisions (Gigerenzer and Goldstein, 1996).

The significance of PMM theory is that it challenges the classical and overly complex view that decisions are the products of elaborate mental calculation, and claims instead that there are simple, parsimonious psychological methods for making inferences that an organism can carry out under real world constraints of limited time and knowledge. Memory search for relevant information is reduced to a minimum (Michon and Pakes, 1995), and rather than integrating what is known to achieve a result, when there are gaps in knowledge other items are substituted for the missing pieces. PMMs perform intelligent guesses about unknown features of the world, based on uncertain indicators (Gigerenzer and Goldstein, 1996). The authors explain that for an organism to reach a conclusion about which of two objects has a higher value, knowledge about a reference class must be examined. The knowledge consists of probability cues. A matrix made up of objects along one axis and cues

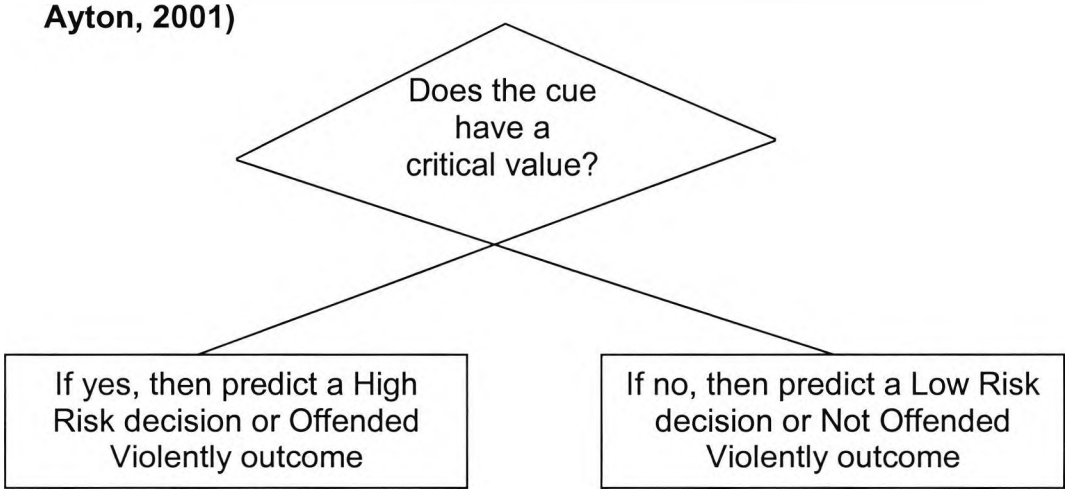
along the other will inevitably have missing entries because the organism will have incomplete knowledge of the world. There will be some things it just does not know. To make a decision the best data is chosen, and the rest ignored. This has led to the development of special types of heuristics.

2.3 Simple heuristics

Some basic procedures in the PMM group are known as the matching heuristic, and the "Take The Best" heuristic (Gigerenzer and Goldstein, 1996).

In the former model, a search is conducted through K of the available cues in rank order of importance, looking for a critical value that indicates a particular decision. If a critical value is found, the search stops, and a decision is based on that information alone. If a critical value is not found, the model searches for the next rank ordered cue. The procedure continues until K cues have been considered. If by this stage no critical value has been found, a decision is made anyway, and no further options are reviewed (Dhami and Ayton, 2001). The matching heuristic for my study, using only one cue, may be illustrated in Figure 3.

FIGURE 3 FLOWCHART OF MATCHING HEURISTIC WHICH SEARCHES FOR JUST ONE CUE (adapted from Exhibit 2, p.156, in Dhami and Ayton, 2001)



The latter is a satisficing process used when inferences must be made quickly about which of two objects scores higher on some criterion, although it can be generalised

to continuous predictors (Gigerenzer, 1996). The underlying principle is that the decision maker will take the best and most easily available piece of information which allows a discriminatory judgement to be made, while ignoring all other data. It assumes a subjective rank order of cues according to their validities, with the highest ranking cue that discriminates between the two alternatives known as the best one (Gigerenzer and Goldstein, 1996).

The process begins with the judge considering whether or not they recognise an object as a predictor of the target variable. The recognition rule states: "If only one of the two objects is recognised, then choose the recognised object. If neither of the two objects is recognised then choose randomly between them. If both of the objects are recognised, then proceed ...," (Gigerenzer and Goldstein, 1996, page 653).

The next stage is to retrieve the value of the highest ranking cue from memory. This will discriminate between the two objects if one has a positive cue value (ie. a higher worth in relation to the target variable) and the other does not. If this distinguishes one from the other, then stop searching for any more information. However, if differentiation is not possible, then consider the next highest ranking sign available from memory, and keep doing so until a choice can be made (Gigerenzer and Goldstein, 1996).

Although the process seems deceptively simple, it can have useful applications in important areas. Breiman et al (1993) described a heuristic that could be used to divide those who had suffered heart attack into high (i.e. those who will not survive at

least 30 more days) or low risk groups. The predictors were ranked in order of merit, each patient allotted a value in relation to these cues, and sequential comparison between the person and physical signs carried out. The first cue was a minimum systolic blood pressure higher than 91mms of mercury over the initial 24-hour observation period. If not, then no other information was considered, and the person was categorised as high risk. If so, then the assessor went on to consider a second cue (ie. if the patient was older than 62 years and 6 months). If not, then the search was stopped and a decision of low risk made. However, if the victim was more than 62 years and 6 months old, then a third and final predictor (sinus tachycardia) was checked, and a classification obtained. The authors reported this simple approach as more accurate than the standard classification method, which measures and integrates 19 variables over the first 24-hour observation period. The practical benefits are obvious. In an emergency, the doctor can quickly and easily obtain information about a small number of variables, and does not need to perform any calculations, because there is no integration of data. This is a characteristic feature of satisficing heuristics which separates them from models used in classical rationality. There is no integration of predictors, no compensation between cues (ie. remaining positive values cannot reverse a decision), and only a limited search is carried out (Gigerenzer, 1996). Yet these fast and frugal psychological principles can lead to decisions as accurately as rational statistical models, sometimes with only one piece of information.

2.4 Rational decision making models

These heuristics embody two fundamental principles of the classical approach to judgement under uncertainty. They use all available information or cue values (ie. complete search), and they combine all this data into a single value (ie. complete integration).

Four of them merit closer examination, and are presented here from the work of Gigerenzer and Goldstein (1996).

One of the simplest is tallying. The number of positive cue values for each object is counted up and the one with the highest score is chosen.

A similar approach is known as weighted tallying. Each cue is given a value based on its perceived importance to the choice under consideration. The figures are summed, and a decision made in favour of the largest total.

Weighted tallying requires more information than simple tallying. In the former, quantitative data about the ecological validity of each cue is necessary. Both these heuristics treat negative and missing information identically, in that they ignore it. Consideration is only given to positive evidence.

In a weighted linear model, cue values are multiplied by their particular ecological validities, and a decision based on the largest total (Michon and Pakes, 1995). This can, however, be a complex task. It is not always clear how to assign relative weights to pieces of information. In the realm of legal decision-making, Wagenaar

et al (1993) found several instances where judges under or over-valued certain bits of evidence in spite of sound empirical findings concerning the value of that material. Although this reflects the differing contributions various cues make to arriving at a judgement, it does not take into account dependencies which may exist between cues.

Multiple regression does, in that it allows for the covariances between predictors. It is commonly seen as an optimal way to integrate diverse bits of information into an estimate. Scientists have thought of this model as simulating the way in which a rational being makes a decision, but there are now challenges to this view.

2.5 Comparison of models using empirical evidence

The research literature on inductive inference has few examples of how well a satisficing heuristic performs in a real-world environment when compared with the more traditional, rational methods (Gigerenzer and Goldstein, 1996). The authors decided to test the accuracy of these two widely differing approaches by seeing how they would cope with a relatively straight forward decision-making task. The problem was to decide which of a series of paired German cities had the larger population, basing the judgement on nine ecological cues which were presented as known (positive) or unknown (negative).

The environment (or reference class) was the set of all German cities with more than 100,000 inhabitants. Population was the target variable. The ecological validity of a cue was the relative frequency with which it correctly predicted the target, defined with respect to the reference class. The nine cues had ecological validities ranging from 0.51 (only slightly better than chance) to 1.0 (absolute certainty). The discrimination rate of a cue was the relative frequency with which it discriminated between any two objects from the reference class.

The larger the ecological validity of a cue, the better the inference. The larger the discrimination rate, the more often a cue can be used to make a decision.

Gigerenzer and Goldstein (1996) used computer simulation to mimic the judgements of 500 people with varying degrees of knowledge about these cities. To model limited recognition knowledge, they imitated people who recognised between zero and 83 cities. To represent limited awareness of cue values, they chose six basic

categories of subjects who knew 0, 10, 20, 50, 72 or 100 per cent of the cues associated with the objects they recognised. There were in total 3,403 inferences made by each "person".

The data base generated for answering the question, "How well did individuals do when using the Take The Best algorithm to decide which city of each pair had more inhabitants", was extensive. For example, the ultimate number of tests conducted was 500 x 6 x 83 x 3,403, or about 858 million.

Gigerenzer and Goldstein (1996) found that the Take The Best heuristic resulted in as many correct decisions as any of the classical, integration models, and more than some. Only slight differences were found between multiple regression, weighted tallying, tallying, and Take The Best. The integration approach which made use of both positive and negative information (ie. the weighted linear model) made considerably fewer correct judgements than Take The Best. All the models did equally well with a complete lack of knowledge, or with complete certainty, but they differed when knowledge was limited. The Take The Best heuristic made as many correct deductions as weighted tallying, and more than the others. It was also the fastest, getting by with less information than the rest, and the authors judged it overall as the highest performer.

This may have been the first time that a satisficing heuristic was shown to produce as many correct decisions about a real-world environment as integration models. The precepts of classical rationality would have led one to expect the latter to do a substantially better job of prediction.

The proven utility of the Take the Best heuristic under conditions of limited or incomplete knowledge, particularly when one considers its simplicity relative to classical, integration models of human judgement, has lead me to wonder what part such procedures might have to play in making predictions about whether or not people referred for risk assessment go on to offend violently.

It is also of interest to see how such a simple rule might perform in terms of forecasting decisions about classifying people into low or high risk of violent offending. A comparison would then be possible between a fast and frugal judgement technique, and the regression model so widely accepted in the psychological literature as representing how rational beings arrive at decisions.

Chapter 3

Methodology

3.1 Problems and aims

Mental health practitioners are regularly asked to make predictions about whether or not patients under their care are likely to behave in a violent fashion (Blumenthal and Lavender, 2000). To date, this decision-making task has been based on clinical or actuarial approaches (Dawes, Faust and Meehl, 1989). In the former, judgement is reached idiosyncratically as a result of the individual decision maker's training and experience, and the steps between problem identification and resolution are not always clearly visible. With actuarial methods the final decision is anchored to objective data, and each stage in reaching it is transparent and able to be replicated. Research findings support the reliability and validity of actuarial techniques over clinical ones (Meehl, 1954; Goldberg, 1965; Dawes, 1971; Dawes, 1979; Lidz, Mulvey, and Gardner, 1993; Grove and Meehl, 1996). However, this is not achieved without cost (in terms of time involved), and anxiety on the part of clinicians that they may be overlooking important information or overly attending to features which on the surface appear powerfully related to future violence. Prediction instruments currently in use such as the PCL-R, the HCR-20, and the VRAG take time to complete, and some of the data they call for is not always available despite extensive file search, meetings with staff members involved in the patient's management, and interviews with significant others (Hare, 1991; Harris et al, 1993; Webster et al, 1997b; Dolan and Doyle, 2000). This exaggerates uncertainty about the result, something already open to controversy.

The prediction of behaviour is a complex task with significant implications for all concerned, which ought to push any sincere scientist-practitioner towards means of doing it as well as the existing knowledge and technology allow. Yet we are not mere super-calculators, capable of integrating a host of forecasting variables, weighting them appropriately, and arriving at a confident decision (Gigerenzer and Goldstein, 1996). Fast and frugal reasoning using probabilistic mental models may not only more closely represent how human decisions are actually made, but might also prove just as efficient (and certainly quicker) than strategies currently in vogue.

This has provided the impetus for the present study, namely, could a fast and frugal reasoning model be applied to the prediction of violent offending, and decisions about which category of risk (i.e. low or high) a patient might pose.

There are 12 research hypotheses (H_1 or alternative hypotheses; Siegel, 1956) in this study. They concern the accuracy with which the three models used can separate the 100 patients in the study cohort into the group which offended violently, the group which did not offend violently, those who were assessed as at high risk of violent offending, and those who were noted as being at low risk of committing violence.

The 12 null hypotheses (H_o; Siegel, 1956) are that none of the three models used will be able to separate patients accurately into the Offended Violently, Not Offended Violently, High Risk, or Low Risk groups any better than chance.

Both research and null hypotheses are shown in Table 15 below:

TABLE FIFTEEN: RESEARCH HYPOTHESES

		OUTCOME		ASSESSORS' DECISIONS	
		Patients who offended violently	Patients who did not Offend violently	Patients assessed as high risk of committing violence	Patients assessed as low risk of committing violence
MODELS USED	Logistic Regression (10 cues)	ONE accurate separation at a better than chance rate	TWO accurate separation at a better than chance rate	THREE accurate separation at a better than chance rate	FOUR accurate separation at a better than chance rate
	Logistic Regression (two most significant cues)	FIVE accurate separation at a better than chance rate	SIX accurate separation at a better than chance rate	SEVEN accurate separation at a better than chance rate	EIGHT accurate separation at a better than chance rate
	Matching Heuristic using the cue with the highest level of predictive accuracy	NINE accurate separation at a better than chance rate	TEN accurate separation at a better than chance rate	ELEVEN accurate separation at a better than chance rate	TWELVE accurate separation at a better than chance rate

3.2 Design of the research

In order to test the null and alternative hypotheses at the same time as developing a fast and frugal reasoning heuristic, a number of cues acknowledged by consultant grade psychiatrists as useful in predicting violence were obtained through postal survey. Details of this process can be seen in Section 3.4, and the result was a data collection matrix shown in Appendix B2. This is made up of the first ten cues from the survey which were also supported by evidence from the literature.

When considering what sample size to choose, attention was paid to the ratio needed between cues and cases. Cooksey (1996) states, "Tabachnick and Fidell (1989) and Hair, Anderson, Tatham, and Black (1992) indicate that for multivariate statistical analyses in general and multiple regression analysis in specific, a ratio of at least 10 cases (cue profiles judged) to every cue utilised is preferred if stable and potentially generalisable regression estimates are desired", (page 123).

Therefore, 100 Kent Forensic Psychiatry Service risk assessment reports written by either psychiatrists or clinical psychologists were examined for these features, and their presence (+), absence (-), or unavailability (?) recorded. Note was also taken of the report writer's risk conclusion at the time (ie. Low or High), and a search carried out on the Police National Computer to determine if the patients concerned had subsequently offended violently (OV) or not (NOV). Files were chosen so that there was at least a two year interval between Kent Forensic Psychiatry Service assessment and computer check, which was done involving the experimenter and one member of Police staff. The information shared was the patient's name, date of

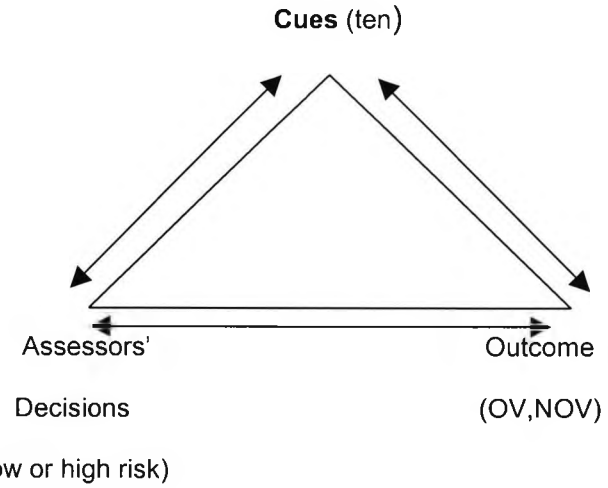
birth, and whether or not they had been registered with the Police over the post-assessment period for a violent offence. The following comparisons were made:

Figure 1 INTERACTIONS BETWEEN STUDY VARIABLES

Adapted from

Figure 13.2, page 238,

In Jackson (1997)



Patients involved in the retrospective file study were chosen after exhaustive examination of the record books kept from Kent Forensic Psychiatry Service referrals allocation meetings. Each week the multi-disciplinary team assembles to discuss requests made by solicitors, the Courts, the prison service, probation staff and medical or psychology colleagues. The emphasis in many of these cases is violence risk assessment. Patients have said or done things which make those charged with their care or supervision feel concerned about the potential for violent offending, and so an expert opinion is sought from our service. Case details are reviewed by the team, attention is given to the time frame involved, and the referral allocated according to an estimate of the best fit between the questions being asked , experience and seniority of staff available, possible conflict of interests, and distance to be covered. Usually two or more people carry out the evaluation after reading the

relevant paperwork, talking to significant others for collateral information, and informing the patient of limits to confidentiality. Interviews can last up to 2½ hours, after which further discussion allows the designated report writer to arrive at a formulation of the presenting problems, reflect on possible treatment or management strategies, and draft a response. The edited draft is sent to the referral agent, and a copy kept in Service files. Verbal feedback is often given to the rest of the team at subsequent meetings, and a system of supervision is in place where people can take problems they encounter as part of this work.

To ensure at least a two year interval between assessment and checking patients' offending behaviour on the Police National Computer, only evaluations carried out prior to January 2000 were considered. An alphabetical list of names was compiled from the referral books where any indication had been given that a violence risk assessment was being asked for. Information about each case in the books was often incomplete, but it was possible to divide patients into definite requests for evaluation of risk, and probables. From the former a cohort of 100 reports was finally selected and read for the cues in Appendix B2.

After the cue data had been collected, contact was made with staff at Maidstone Police Headquarters. A good working relationship exists between the Kent Forensic Psychiatry Service and Police, locally but also nationally. Permission was obtained for the experimenter to access computer records under the direct supervision of police staff. The information looked for was a match between name, date of birth, and an official entry that this person had been cautioned for, charged with, or convicted of a violent offence since January 2000. The former categories were

included because the legal process can be slow, and relying only on the latter might provide too restrictive a sample. Some might say that adopting this rigorous a criterion biases the study, because it does not take account of allegations of violence, or suggestions from collateral informants that the patient is behaving aggressively but has not come to the attention of the authorities. However, the purpose of this piece of research is to move away from subjective or vague assessments of risk, and so the generally accepted legal standards of presumed innocence and proof beyond reasonable doubt that violence has occurred have been used as boundary measures.

This data may be seen in the last two columns of Appendix B2.

3.3 Ethical considerations

In a study of this nature, where sensitive and personal information about people is read by a third party, and identifying features such as name and date of birth are used in conjunction with a statutory agency (ie the police) to determine if offences have occurred, consideration must be given to how this might be done safely.

Application was made to various individuals and committees in accordance with the policies of the West Kent NHS & Social Care Trust regarding this project. These requests will be summarised here, and the detailed information provided can be found in a series of attachments beginning with Appendix B6.

On 18th February 2002 the application form for submission of a research proposal to the West Kent Research Ethics Committees was completed and posted. It contained a summary of the project, details of what would be involved, the principle investigator's curriculum vitae, and may be seen in its entirety in Appendix B6.

Later in February 2002, an undated letter was received from Ms Julie Knowles (senior administrator, Ethics Committees) acknowledging receipt of the application form, and suggesting amendments before appearing at a committee meeting on 27th March 2002. The Chairman required the summary of the project to be written in clear language that a lay person could understand, for me to inform Dr Claude Pendaries of my intended project, and a letter from my university supervisor confirming that he supported my research. These requirements were met, and Ms Knowles' letter may be seen as Appendix B7.

A letter dated 27th February 2002 was received from Dr Claude Pendaries (Director of Corporate Affairs, Lead Research & Development Officer). It stated that I must notify him of my intention to undertake a research project, and that after approval had been given by the Research Ethics Committee I must contact him again to seek approval to begin. This letter may be seen in Appendix B8.

On 5th March 2002 I wrote to Ms Knowles enclosing the amendments requested, and confirming my intention to appear before the Research Ethics Committee meeting on 27th March 2002. This letter forms Appendix B9.

On 5th March 2002 I also wrote to Dr Pendaries with the project details he required, and this letter may be seen in Appendix B10. In it I spelt out to him that several of the reports I wanted to review in the retrospective file study would have been written by me, others I may have co-authored, still more I would have sat in on the actual interview with the patient, and virtually all of the rest would have been discussed in my presence at either the referrals allocation meeting where the patient was first considered, or subsequent meetings when the results of evaluations were shared with senior clinicians. I also explained to him that after discussions with police staff I had been made aware that the information I sought was actually a matter of public record. I could have gone to the offices of the Kent newspapers and checked back through their Court pages for the last two years, looking for matches between names on my list and people they have reported on. This would of course have been enormously time-consuming, and the only reason I sought to use police resources (ie their computer database) in the first place was to condense the data gathering process.

Ms Knowles wrote to me (a letter dated 15th March 2002) confirming that I would be interviewed by the Research Ethics Committee on 27th March 2002. Proof of this may be seen in Appendix B11.

Dr J B Symes (Chairman – Dartford & Gravesham Research Ethics Committee) wrote on 5th April 2002 detailing points raised at the 27th March 2002 meeting. The Committee required a letter from the police acknowledging that I could access their database, the inclusion of cautions to patients by the police as evidence of offending, a written guarantee that data gathered would be stored securely/seen only by me/ destroyed confidentially once the project was completed, and Dr Peter Ayton's written confirmation that he supported the project. The last paragraph states, "The Committee agreed that there was no objection on ethical grounds but before formal approval can be given the Committee require sight of the above information". Appendix B12 contains Dr Symes' letter.

A letter of support from Dr Peter Ayton dated 15th April 2002 may be seen in Appendix B13.

On 29th May 2002, Dr Symes wrote to me requesting an update on the position of my study, in particular asking what had been done about the four points raised by the Committee at the 27th March 2002 meeting. This letter may be seen in Appendix B14.

On 18th June 2002 I wrote back to Dr Symes, responding to the four matters raised, and enclosing supportive documents from the people concerned. This letter forms Appendix B15.

A letter dated 18th June 2002 from Detective Sergeant Christopher Tomlin to Dr Symes gives evidence of his support for, and level of involvement in, the research project. It is labelled as Appendix B16.

On 5th July 2002, Dr Symes wrote to me acknowledging that the requirements of the Research Ethics Committee had been met, and that there was no objection on ethical grounds for the study going ahead. That letter may be seen in Appendix B17.

On 17th July 2002, Dr Pendaries wrote to me stating that my project had been approved by the Trust, and formally registered into the Trust's research database. The letter may be seen in Appendix B18.

The final piece of correspondence received about ethical approval of this study came from Dr Symes (dated 5th August 2002), and confirmed that the Committee ratified his Chairman's action in approving my project as detailed in his letter of 5th July 2002. Appendix B19 has the August 5th letter in it.

3.4 Postal survey of cues selected by consultant psychiatrists for predicting violence.

When this project was at the planning stage, it was decided that a list of statements would be abstracted from the Hare Psychopathy Checklist – Revised (Hare, 1991), and sent to consultant psychiatrists with a request that each one be rated on a ten-point scale for usefulness when attempting to predict if a patient would behave violently. Discussions with psychiatric colleagues in our directorate led me to believe that the time it would take each doctor surveyed to complete the task may well reduce the number of forms returned, so this approach was abandoned. The next option was to select cues to future violent conduct from the literature (e.g. Bonta, Hanson, and Law, 1998), and use them in a retrospective file study to see if they had any predictive power. After considering this with my supervisor (Dr Peter Ayton) it was agreed to return to the concept of a postal survey, but with a format which took very little time to complete. Respondents were sent a covering letter of one page, with another sheet attached where they were asked to "Please write in the space below the cues you look for when evaluating the potential for a patient of yours to behave violently in the future", (see Appendix B1). A postage-paid envelope was included with each letter.

The membership list of the Royal College of Psychiatrists (2000) was checked for the names of 100 consultant grade doctors, and the posting of forms took place between June 4th and 15th of 2001. By July 12th 2001, 43% of them had been returned. Of these, six were not able to be included in the study. The responses were;

6 "I find it impossible to predict violence and try to but have not been successful yet, sorry".

30 "No longer at this address."

45 "I cannot complete this because it gives no time frame for the violent behaviour. Is this immediate threat or long-term risk following discharge?"

61 "I'm afraid I can't answer in this format. The question is far too wide to allow meaningful responses. Sorry."

93 "Sorry – too many forms this week".

99 "This doctor is now retired."

The small number of unusable replies, and the lack of any consistent critical theme among them, led me to conclude that the survey form was reasonable and unambiguous. This view was supported by the 37 consultants who were able to identify cues they looked for when attempting to predict violence. This cohort cited a total of 83 factors they thought important when making such a decision. There was considerable overlap of choices, with many items selected by a number of consultants. In descending order of popularity those chosen by five or more may be seen in Table One below.

TABLE ONE: CUE FREQUENCIES FROM POSTA\L SURVEY RESPONDENTS

Item	Tally
History of past violence	30
Substance or alcohol misuse history	27
Threats or impulses to violence	18
Active symptoms of mental illness	16
Delusions of persecution	12
Non-compliance with medication	10
Use of weapons	9
Mood disorder	9
Threat/Control over-ride permitting phenomena	9
Availability of victim	8
Disorganised social circumstances	8
Attitude at interview	8
Fantasies of violence	7
Long-term relationship problems	7
Irritability	6
Current stressors	6
Mental illness associated with past violence	5
History of socialised violence	5
Paranoid, borderline, or antisocial personality disorder	5

A history of aggressive behaviour has been identified as a principal indicator of future violence in a number of studies (e.g. Swanson, 1993; Litwack, 1994; Morrison, 1994; Torrey, 1994; McNiel and Binder, 1995; Bjorkly, 1997; Gilders, 1997; Monahan et al, 2000). Past behaviour is one of the most stable predictors of

future actions, and clinicians making judgements about risk tend to view the person's history of violence as of paramount importance (McNiel and Binder, 1995). Mossman (1994) found that prediction based solely on a patient's previous history of violence was nearly as accurate as other forms of actuarially produced decisions. Bonta, Hanson and Law (1998) reported that criminal history variables (including violent background) were much better forecasters of future offending than were clinical features. The fact that this cue was the one most heavily endorsed by participants in the postal survey lends some support to the belief that consultants are looking for appropriate indicators when attempting to predict violence.

Research on the contribution of substance abuse to criminal behaviour indicates that the use of alcohol and drugs often precedes violent offending (Blumenthal and Lavender, 2000). Studies have indicated overwhelmingly that the majority of violent crimes (e.g. assault, sexual assault, family violence, murder) are committed by individuals who are intoxicated with alcohol (Fitch and Papantonio, 1983; Pihl and Peterson, 1993; Beck, 1994; Bergman and Brismar, 1994; Modestin, Berger and Ammann, 1996; Bonta, Hanson and Law, 1998; Monahan et al, 2000). Among 26 studies involving 9304 cases Murdoch, Pihl and Ross (1990) found that 62% of violent offenders were drinking at the time of the offence. Again it would seem that the participants in this survey are operating within the doctrine of evidence based practice when selecting substance or alcohol misuse as an important indicator of the likelihood of future violent offending.

The evidence is not so strong, however, for the 3rd highest ranked cue (i.e. threats or impulses to violence). One paper by Grisso et al (2000) used data from the

McArthur Risk Assessment Study, a project involving the follow-up of 1136 discharged psychiatric patients. The prevalence of self-reported violent thoughts among hospitalised psychiatric patients was compared with a control group in the community. Subjects were followed up after discharge and the persistence of violent thoughts investigated. The reporting of violent thoughts was significantly related to being violent within 20 weeks of discharge for non-white patients, for those without major mental disorder but with a diagnosis of substance abuse, for patients with high symptom severity during admission, and for those who still complained of violent thoughts after discharge from hospital. These findings were seen as having more utility in assisting clinicians to identify relevant areas of enquiry during risk assessment, rather then serving as indicators of future violence. Perhaps consultants in the postal survey have been drawn to the face validity of such threats rather than being aware of how little they contribute on their own to the prediction of violence.

There is a growing recognition of the limited utility of diagnosis as a risk factor, and an increasing focus on active symptoms of severe mental disorder instead (Blumenthal and Lavender, 2000; Bjorkly, 1997). To quote Monahan (2000), "Active symptoms are probably more important as a risk factor than is simply the presence of an identifiable disorder", (page 315). Krakowski et al (1986) concluded that the frequency of violent incidents tends to reflect the course of the acute psychosis and to be positively correlated with the severity of psychotic symptoms. In their review of a series of studies, Wessely and Taylor (1991) concluded that between one and two thirds of homicidal violence committed by individuals with schizophrenia could be attributed to abnormal mental state. It seems, therefore, that survey respondents

have some support in choosing to place active symptoms of mental illness 4th in their ranking.

At 5th place were delusions of persecution. These are defined by Taylor et al (1994) as "based on an absolute conviction of the truth of a proposition which is idiosyncratic, incorrigible, ego-involved, and often pre-occupying", (page 163). Violence by the mentally disordered often appears to be a rational response to an irrational belief (Link and Stueve, 1994). In a study involving 1740 special hospital patients, Taylor et al (1998) found that 75% of those with a diagnosis of functional psychosis were recorded as having been motivated to offend by their delusions. Shore et al (1989) followed up a group of mentally disordered individuals arrested near the White House. For those with no history of previous violence, there was a significant relationship between delusions of persecution and future violence. Risk of aggression increases if paranoid delusions appear in combination with auditory command hallucinations that instruct the patient to render other people harmless as part of a "self-defence plan" (Taylor, 1985; Taylor and Gunn, 1984). It seems therefore reasonable to take delusions seriously when attempting to predict violence.

Respondents ranked non-compliance with psychotropic medication regimes as 6th in importance. This has been identified as a useful predictor of dangerousness in those with a psychotic disorder (Torrey, 1994; Monahan et al, 2000). A review by Howlett (1998) estimated that 70% of psychiatric patients discharged from hospital were likely to stop taking their medication within a period of two years. He examined independent inquiry reports into homicides by the mentally ill. Out of 35 cases, 20

(or 57%) were found to have non-compliance with medication as a major contributory factor in the breakdown of care prior to the killing.

Use of, facility with, and collecting weapons have been accepted as worthwhile predictors of potential for future violence (Bonta et al, 1998). This cluster was rated 7th by consultants responding to the postal survey.

These doctors viewed mood disorder as important in risk assessment, but the literature does not provide support for such a belief. Bonta, Hanson, and Law (1998) concluded from their meta-analysis that mood disorders were unrelated to violent recidivism for both mentally ill and non-disordered offenders. Similarly, Monahan et al (2000) reported non-significant correlations between mania, depression, and violence. On this basis there would seem to be no good grounds for using this cue in data collection for the retrospective file study.

By contrast, there is some support for including the 9th ranked cues in such an exercise. Symptoms associated with threat (the feeling that others wish to harm you) or the over-riding of personal controls (the belief that your mind is dominated by forces outside your control, or that thoughts are being inserted into your mind which are not your own) have been found to account for violence by individuals with severe mental disorder (Blumenthal and Lavender, 2000). Participants who reported threat/control over-ride symptoms in the Swanson et al (1996) study were twice as likely as those reporting other psychotic symptoms to engage in violent behaviour, and about five times as likely as those with no mental disorder. Similarly, Monahan

et al (2000) found a statistically significant relationship between threat/control-override symptoms and violence.

The availability of an identified target (or potential victim) for hostility was ranked 10th by postal survey respondents. There is support for this view in the literature. Blumenthal and Lavender (2000) observed that the majority of violence by individuals with severe mental disorder is against relatives (most frequently mothers), and that few strangers are attacked. Taylor and Gunn (1999) studied changes in Home Office figures of homicide by people with mental illness convicted of manslaughter on the grounds of diminished responsibility in England and Wales over a 38-year period. Attacks on strangers were the exception. In his view of a series of studies, Bjorkly (1997) concluded that close family members, particularly parents, were at higher risk of becoming victims of mentally ill patients' aggression than other people. Monahan et al (2000) noted statistically significant correlations between violence and the identification of a specific target for aggression.

Eleventh was disorganised social circumstances. The neighbourhood in which an individual resides has been identified as a highly significant predictor of the likelihood of violence (Blumenthal and Lavender, 2000). A dysfunctional family background was noted as a powerful predictor of violent recidivism by Bonta et al (1998).

The final cue among the top 12 was the patient's attitude at interview. Eight consultants chose this item as influencing their decisions about future risk (i.e. 8 out of 37, or 22%), but no support for its accuracy could be found in the research literature, so it was not used.

After comparing survey results with published evidence, the original list of cues was reduced to 10. Mood disorder and attitude at interview were omitted on the grounds that there was insufficient proof to justify including them. The final data collection format for the retrospective file study may be seen in Appendix B2.

3.5 Data collection cues defined

These are as follows:

3.5.1 History of violence

Any aggressive act against another person for which the patient was cautioned by the Police, arrested, convicted in Court, or admitted to a psychiatric hospital under a section of the Mental Health Act 1983 and its subsequent revisions. This must have been verified by a collateral source (such as the referral agent, another professional contact, or a relative).

3.5.2 Substance or alcohol misuse history

The taking of illicit drugs or consumption of enough alcohol to make the person feel intoxicated at a greater frequency than is usually reported, and which exceeds officially accepted WHO levels for safe use (i.e. up to 28 units per week). This must be of sufficient severity to produce disturbances in level of consciousness, cognition, perception, affect, or behaviour that are of clinical importance (ICD-10, 1993).

3.5.3 Threats or impulses to violence

Subjective reports at interview that the person has seriously contemplated becoming physically aggressive to somebody else, or written comments made by the referral agent to this effect.

3.5.4 Active symptoms of mental illness.

Acknowledgement by the individual at interview of diagnostic signs accepted by the interviewer as being consistent with ICD-10 categories of mental illness.

3.5.5 Delusions of persecution

Acknowledgement by the individual at interview of a longstanding but irrational belief, held with absolute conviction, that some person, group of people, or agency means them harm.

3.5.6 Non-compliance with medication

Admission at interview or indication by the referral agent that the person has been prescribed psychotropic medication by a registered medical practitioner for the treatment of a psychiatric disorder, but that they have taken it inconsistently, or have stopped altogether.

3.5.7 Use of weapons

Reference by the referral agent or the patient themselves to familiarity with, interest in, collections of, or access to potentially dangerous implements unlikely to be readily available in most homes (e.g. firearms, swords, archery equipment, crossbows, throwing knives, rice flails). This can, however, also include the use of household items or available objects to threaten or hurt somebody.

3.5.8 Threat/control – over ride permitting phenomena

Report by the patient or referral agent of symptoms associated with threat (the feeling that others wish to cause harm) or the over-riding of personal control (a belief that one's mind is dominated by forces outside of the individual's control, or that thoughts are being inserted into the person's mind which are not their own).

3.5.9 Availability of victim

Mention by the patient or referral agent of a specific named person to whom there is access as a possible target for future aggression.

3.5.10 Disorganised social circumstances

Mention by the patient (or referral agent) of current financial pressures, association with people engaging in alcohol misuse or illicit substance use, criminal activity, dissatisfaction with partner, or uncertainty about accommodation arrangements.

Chapter 4

Results

4.1 Sample characteristics

The research cohort was made up of 90 males and 10 females. The youngest person at date of interview was 17 years and four months old, while the eldest was 66 years and five months.

Of the 1300 cells in Appendix B3 comprising the data matrix obtained from risk assessment documents about patients in this study, only eight contained missing or incomplete entries (as denoted by 2). This works out at 0.62% of the total.

The earliest appointment date was 16.1.96 for patient number 52, and the latest was 23.12.99 for patient number 23. When one takes into account the date on which the Police National Computer Data Base was accessed to evaluate outcome (i.e. 23.5.02) this means the period of follow up, during which people had the opportunity to offend violently, ranged from 2 years 5 months to 6 years 4 months.

The actual intervals between assessment and violent offending may be seen in Table Two below.

TABLE TWO: TIME INTERVAL BETWEEN ASSESSMENT AND VIOLENT OFFENDING FOR THE TWELVE VIOLENT OFFENDERS IN THE SAMPLE

Number of months between assessment and offence	Assessment Date	Offence Date	OV Patient Number
14	4.12.78	10.2.00	7
21	14.9.99	7.6.01	8
24	16.3.98	18.3.00	10
27	15.2.99	22.5.01	21
2	12.7.99	7.9.99	24
29	8.4.99	18.9.01	27
14	22.11.99	16.1.01	29
43	23.7.97	25.2.01	32
29	4.6.99	15.11.01	54
45	25.2.97	23.11.00	55
63	30.4.96	19.8.01	74
21	3.11.98	20.8.00	94
Total : 332 months			
Average : 28 months			

Although the range of variation was between 2 and 63 months, the average gap was 28 months.

Distinctions made amongst patients in the sample were those rated at interview as posing a high risk of behaving violently (HR), those considered to be low risk (LR), and subsequently who out of both groups went on to offend violently by the time of follow-up (OV), and those who had not (NOV).

Frequencies and percentages for these categories may be seen in Table Three below.

TABLE THREE : FREQUENCY OF PREDICTIVE CUES AND THEIR RELATION TO DECISIONS AND OUTCOMES

Predictor Number		1	2	3	4	5	6	7	8	9	10			
Predictor		History of violence	Substance or alcohol misuse history	Threats or impulses to violence	Active symptoms of mental illness	Delusions of perse-cution	Non-compliance with medication	Use of weapons	Threat/ Control-Override permitting phenomena	Avail-ability of victim	Disor-ganised Social Circum-stances	Risk Asst L H	O V	N O V
Whole Sample (100)		72	58	81	28	25	18	43	15	47	40	L54 H46	12	88
High Risk Group (46)	N	39	29	41	16	15	12	21	12	22	19	L. 0/H.46	9	37
	%	85	63	89	35	33	26	46	26	49	41	L.0/H.100	20	80
Low Risk Group (54)	N	33	29	40	12	10	6	22	3	25	21	L.54/H.0	3	51
	%	61	54	74	22	18	11	41	5	46	39	L.100/H.0	6	94
OV Group (12)	N	11	8	12	1	1	1	8	0	6	8	L.3/H.9	12	0
	%	92	67	100	8	8	8	67	0	50	67	L.25/H.75	100	0
NOV Group (88)	N	61	50	69	27	24	17	35	15	41	32	L.51/H.37	0	88
	%	69	57	78	31	27	19	40	17	47	36	L.58/H.42	0	100

From the sample of 100 patients, 46 were assessed as HR, and 54 as LR. At follow-up, 88 had not offended violently, while 12 had.

4.2 Sample members who had offended over the follow-up period

One of the offenders was female (participant number 10), while the other 11 were males.

Ages at the end of the follow-up period (ie 23/5/02) ranged from 20 to 33 years, with the average being 27 years and eight months.

Assessors had rated nine of these people as high risk of offending violently, and three as low risk. The offences may be seen in Table Four below.

TABLE FOUR: RISK ASSESSMENT AND OFFENCE FOR THE TWELVE SAMPLE MEMBERS WHO COMMITTED VIOLENT CRIMES

Patient No.	Risk Assessment	Offence
7	L	Using threatening, abusive, insulting words or behaviour
8	H	Gross indecency with a child. Permitting the taking of indecent photographs, and possessing indecent photographs.
10	L	Pursuing a course of conduct which amounts to harassment.
21	H	Possessing explosives for unlawful purposes.
24	H	Common assault.
27	H	Indecent assault on a female under 16 years of age
29	L	Aggravated bodily harm.
32	H	Kidnapping, robbery and grievous bodily harm with intent.
54	H	Aggravated bodily harm and affray.
55	H	Grievous bodily harm with intent. Threatening words or behaviour with intent to cause fear or concern (x2)
74	H	Assault on a constable.
94	H	Assault occasioning actual bodily harm and threats to kill.

Raw data for the Offended Violently group can be seen in Appendix B4. This shows the presence (+) or absence (-) of the 10 cues under investigation with regard to their utility in predicting outcome (Offended Violently or Not Offended Violently) and assessors' decisions about risk (Low Risk or High Risk).

4.3 Rationale for selecting statistical analyses used

The purpose of this research project was to see if a small number of cues considered by assessors predicting risk level could be used to separate accurately members of the study cohort into those who offended violently, those who did not offend violently, those deemed at high risk of violent offending, and those thought to pose a low risk of such behaviour.

In the first instance, correlations between cues, decisions, and outcome were calculated, and those which were statistically significant may be seen in Table Five. All the correlations calculated may be seen in Appendix B5.

A chi-square test was done to examine the strength of association between prediction and outcome (see Section 4.5). The result was a chi-square of 4.617 ($p=0.032$, $df=1$). This figure is statistically significant, and led on to further examination of just how strong the link between prediction of violence and outcome was.

Logistic regression was used for this. It is the appropriate method to select when dependent variables are categorical (ie. membership of the offended violently, not offended violently, high risk, or low risk groups). There is no overlap between these categories. A patient is in one of the outcome or decision groups or the other, but cannot be in both. Logistic regression requires fewer assumptions in theory, and is statistically robust (Lea, 1997). The analysis was done using all ten cues (independent variables) to see how effective this combination was in allocating patients correctly into the offended violently, not offended violently, high risk, and low risk groups. The results may be seen in Sections 4.6 and 4.7.

In an attempt to see how much redundancy there might be amongst the independent variables with reference to their effectiveness in correctly allocating patients to the relevant groups, logistic regression was carried out using only the two most significant cues (ie. disorganised social circumstances and use of weapons for outcome; history of violence and non-compliance with psychotropic medication for assessors' decisions). The results may be seen in Sections 4.8 and 4.9.

Using two predictors instead of ten is a faster, more frugal approach, and in keeping with the purpose of the study. It set the scene for the final model under consideration (ie. the matching heuristic). Here each cue was considered on its own in relation to accuracy rates for predicting membership of the four dependent variable groupings (ie. OV, NOV, LR, and HR).

To predict outcomes, accuracy rates were the sum of cue present and OV group membership with cue absent and NOV group membership. To predict decisions, accuracy rates were the sum of cue present and HR decision with cue absent and LR decision (Dhami and Ayton, 2001). The results are summarised in Table Thirteen.

4.4 Correlation between cues

Easily defined, readily observable cues to predict violent offending were selected after postal survey of consultant psychiatrists, and these can now be compared with each other, with level of risk assessment (i.e. High or Low), and with outcome at follow up (i.e. Offended Violently or Not Offended Violently).

The data matrix comprising Appendix B5 was produced. This is made up of 13 cells x 13 cells. The 10 cues to violence prediction used in the retrospective file study had three extra factors added; risk assessment (Low or High), offended violently or not (OV or NOV), and Psyagg. The latter was an equally weighted aggregate made up of the four cues which had a psychiatric component (i.e. active symptoms of mental illness, delusions of persecution, non-compliance with a psychotropic medication regime, and threat/control-override permitting phenomena). This was done to see if such a combination had more power to predict decisions and outcomes than the individual cues themselves.

Frequency counts were carried out on the data in Appendix B3 (the matrix of 1300 cells obtained from risk assessment documents about patients in this study). Pearson correlation coefficients were calculated to test the strength of association between factors, and these figures may be seen in Appendix B5. In total, 27 correlations achieved statistical significance, eight at the 5% level, and 19 at the 1% level. These are presented in Table Five below.

TABLE FIVE: STATISTICALLY SIGNIFICANT CORRELATIONS BETWEEN CUES, DECISIONS, AND OUTCOME

Cues	History of violence	Threats or impulses to violence	Active symptoms of mental illness	Delusions of persecution	Non-compli-ance with medica-tion	Threat/control - override permitting phenomena	Avail-ablity of victim	Dis-organised social circums-tances	Risk assess-ment
Active symptoms of mental illness		.310**							
Delusions of persecution		.287**	.576**						
Non-compliance with medication		.230*	.634**	.449**					
Use of weapons	.260**	.206*							
Threat/ control – override permitting phenomena	.203*	.206*	.423**	.338**	.275**				
Availability of victim		.322**	.211*						
Risk assessment	.275**	.201*			.286**	.275**			
Offended violently								.209*	
Psychiatric aggregate		.335**	.864**	.772**	.764**	.625**	.230*		.344**

** Correlation is significant at the 0.01 level

* Correlation is significant at the 0.05 level

The next section will consider the strength of association between prediction of violence and outcome.

4.5 Risk assessment compared to outcome

In the prediction of violence there are two ways of being wrong and two ways of being right. A forecast of not dangerous (ND) may be confirmed at follow-up as a true negative (TN). Alternatively, a judgement that an individual will act dangerously (D) may prove correct, yielding a true positive (TP). These combinations are important for measuring accuracy of prediction, and may be seen along with other pairings in Table Six below.

TABLE SIX : VIOLENCE PREDICTION CONTINGENCIES

		PREDICTION: ND	PREDICTION: D
OUTCOME	ND	TN Predicted not dangerous and shown at follow-up to be not dangerous	FP Predicted to be dangerous, but at follow-up shown to be not dangerous
	D	FN Predicted to be not dangerous, but at follow-up found to be dangerous	TP Predicted to be dangerous, and at follow-up found to be dangerous

Adapting this approach in the present study allows the production of Table Seven.

TABLE SEVEN: PREDICTION VERSUS OUTCOME IN SAMPLE

		PREDICTION: LR	PREDICTION: HR
OUTCOME	NOV	TN 51	FP 37
	OV	FN 3	TP 9

A chi-square test to examine the strength of association between prediction and outcome was carried out to produce Table Eight.

TABLE EIGHT: EXPECTED VERSUS OBSERVED SAMPLE FIGURES

		PREDICTION		
		LR	HR	TOTALS
OUTCOME	NOV 54	TN 51 47.52	FP 37 40.48	88
	OV 46	FN 3 6.48	TP 9 5.52	12
	100	54	46	100

Chi-squared equals 4.617 (p= 0.032, df=1).

This result is statistically significant, and indicates there is an association between prediction of violence and outcome. Just how strong this link is will be shown in the following sections.

4.6 Logistic regression analysis using all ten cues to predict outcome

Comparing the ten predictor variables with violent outcome revealed a surprising absence of association. Only disorganised social circumstances and violent offending were significantly correlated. Consideration turned to a logistic regression analysis of the study data in an attempt to predict which of the 100 patients in this cohort could be correctly allocated to the relevant groups.

In the first instance, predictor (or independent) variables were all ten cues identified as likely to have relevance when estimating the possibility of violence, and the criterion (or dependent) variable was whether or not a person was correctly classified into the offended violently group.

The logistic regression technique was chosen because the criterion variable was dichotomous (ie OV or NOV) rather than measured on a continuous scale (Brace, Kemp, and Snelgar, 2000; Cooksey, 1996).

The SPSS was used on all 100 cases to produce a logistic regression model with the following characteristics;

TABLE NINE : LOGISTIC REGRESSION MODEL SHOWING OBSERVED VERSUS PREDICTED VIOLENT OFFENDING USING ALL TEN CUES

		Predicted				
		OV				Percentage
Observed		No		Yes		Correct
OV	No	TN	88	FP	0	100.0
	Yes	FN	12	TP	0	.0
Overall Percentage						88.0

Overall the model has an 88% correct classification rate. It predicted that 88 patients would not offend violently, and this matched up with what was actually observed in the study.

Twelve patients were mis-classified by the model, which predicted that these people would also appear in the NOV group, when in fact the study cohort contained twelve individuals who were observed to have committed violent offences.

The model could only make negative predictions, in other words that all 100 people would appear in the NOV group. None of the independent variables used achieved sufficient levels of statistical significance to allow the model to distinguish between those who offended violently and those who did not.

A full account of this particular phase of the data analysis may be seen in Appendix B20.

4.7 Logistic regression analysis using all ten cues to predict assessors' decisions

The model was tested to see how closely it would approximate divisions made by assessors about whether patients were at low risk of offending violently, or at high risk.

Predictor (or independent) variables were all ten cues identified as likely to have relevance when estimating the possibility of violence, and the criterion (or dependent) variable was whether or not a person was correctly classified into the high or low risk groups.

The SPSS was used on all 100 cases to produce a logistic regression model with the following characteristics:

TABLE TEN : LOGISTIC REGRESSION MODEL USING ALL TEN CUES TO PREDICT ASSESSORS' DECISIONS

		Predicted				
		Decision				Percentage
Observed		Low		High		Correct
Decision	Low	TN	21	FP	33	38.9
	High	FN	7	TP	39	84.8
Overall Percentage						60.0

The model has a 60% correct classification rate, made up of 21 patients who were predicted to be low risk and who were actually classified as such by assessors, and 39 patients who were predicted to be high risk and who were seen this way by assessors.

Forty patients were mis-classified by the model, which predicted that seven would be categorised as low risk (false negatives) when assessors thought these people more closely matched the criteria for a high risk judgement, and 33 who were predicted high risk (false positives) when assessors considered them to be low risk.

A full account of this particular aspect of the data analysis may be seen in Appendix B21.

4.8 Logistic regression analysis using the two most significant cues to predict outcome

For outcome, the two most significant cues were disorganised social circumstances (r=0.209), and use of weapons (r=0.185). This can be seen in Appendix B5.

Putting these cues into a logistic regression model allows us to evaluate its performance against the earlier model using all ten predictors.

Predictor (or independent) variables were disorganised social circumstances and use of weapons, while the criterion (or dependent) variable was whether or not a person was correctly classified into the offended violently group.

The SPSS was used on all 100 cases to produce a logistic regression model with the following characteristics;

TABLE ELEVEN : LOGISTIC REGRESSION MODEL SHOWING OBSERVED VERSUS PREDICTED VIOLENT OFFENDING USING THE TWO MOST SIGNIFICANT CUES

Observed		Predicted				
		OV				Percentage Correct
		No		Yes		
OV	No	TN	88	FP	0	100.0
	Yes	FN	12	TP	0	.0
Overall Percentage						88.0

The model could only make negative predictions, in other words that all 100 people would appear in the NOV group. Neither of the independent variables used (alone or in combination) achieved sufficient levels of statistical significance to allow the model to distinguish between those who offended violently and those who did not.

Overall, the accuracy rate for the model was 88%, as this many people did not commit violent offences.

However, twelve patients were mis-classified by the model, which predicted that these people too would appear in the NOV group, when in fact the study cohort contained 12 individuals who were observed to have committed violent offences.

The failure of the two-cue model and the one using all ten cues to differentiate OV and NOV patients is not surprising, as their respective overall levels of statistical significance are p=0.065 and p=0.079 (see Appendices B22 and B20).

A full account of this particular phase of the data analysis may be seen in Appendix B22.

4.9 Logistic regression analysis using the two most significant cues to predict assessors' decisions

For decisions of low or high risk, the two most significantly correlated cues were history of violence (r=0.275), and non-compliance with psychotropic medication (r=0.286). Both were significant at the 0.01 level (see Appendix B5).

Putting these cues into a logistic regression model allows us to evaluate its performance against the earlier model using all ten predictors. At the same time it might show what level of redundancy there could be amongst the other eight cues.

Predictor (or independent) variables were history of violence and non-compliance with psychotropic medication, while the criterion (or dependent) variable was whether or not a person was correctly classified into the high or low risk groups.

The SPSS was used on all 100 cases to produce a logistic regression model with the following characteristics;

TABLE TWELVE : DECISION PREDICTION USING ONLY THE TWO MOST SIGNIFICANT CUES

		Predicted				
		Decision				Percentage
Observed		Low		High		Correct
Decision	Low	TN	21	FP	33	38.9
	High	FN	7	TP	39	84.8
Overall Percentage						60.0

The model has a 60% correct classification rate, made up of 21 patients who were predicted to be low risk and who were actually classified as such by assessors, and 39 patients who were predicted to be high risk and who were seen this way by assessors.

Forty patients were mis-classified by the model, which predicted that seven would be categorised as low risk (false negatives) when assessors thought these people more closely matched the criteria for a high risk judgement, and 33 who were predicted high risk (false positives) when assessors considered them to be low risk.

The identical results achieved by this two-cue model and the one using all ten cues, suggest that the other eight predictors have nothing to add in terms of enhancing predictive accuracy, and are in effect redundant.

A full account of this particular phase of the data analysis may be seen in Appendix B23.

4.10 Outcome and decision prediction using the matching heuristic

The matching heuristic was calculated for each cue to discover how efficiently outcomes and decisions could be predicted. Appendix B3 data (the matrix of 1300 cells obtained from risk assessment documents about patients in this study) was used in two equations. To predict outcomes, accuracy rates were the sum of cue present and OV group membership with cue absent and NOV group membership. To predict decisions, accuracy rates were the sum of cue present and HR decision with cue absent and LR decision. The results may be seen in Table Thirteen below.

TABLE THIRTEEN: MATCHING HEURISTIC PREDICTING OUTCOMES AND DECISIONS

	History Of violence	Substance or alcohol misuse history	Threats or impulses to violence	Active symptoms of mental Illness	Delusions of perse-cution	Non-com-pliance with medi-cation	Use of wea-pons	Threat/ control override permitting pheno-mena	Availa-bility of victim	Dis-organised social circum-stances
Accuracy rates for predicting outcomes (i.e. OV or NOV) in percen-tages	38	44	31	62	65	71	61	72	53	60
Accuracy rates for predicting decisions (i.e. HR or LR) in percen-tages	60	52	55	58	59	59	53	62	51	51

For outcomes, using only one cue in a matching heuristic produced levels of predictive accuracy ranging from 31% (for threats or impulses to violence) to 72% (for threat/control-override permitting phenomena).

For panel risk assessments the variation was less pronounced, and using only one cue the matching heuristic produced levels of predictive accuracy ranging from 51% (for both availability of victim and disorganised social circumstances) to 62% (for threat/control override permitting phenomena). The meaning and clinical utility of these figures will be considered in detail in the discussion section of this report.

4.11 Summary: accuracy of models with regard to separating patients in terms of outcome and decision

The three models generated (ie. logistic regression using ten cues, logistic regression using the two most significant cues, and the matching heuristic) were tested for their utility in forecasting outcome (OV or NOV group membership), and decisions (HR or LR group membership).

The results are summarised in Table Fourteen below;

TABLE FOURTEEN : ACCURACY OF MODELS IN SEPARATING PATIENTS WITH REGARD TO OUTCOME AND ASSESSORS' DECISIONS

		OUTCOME		ASSESSORS' DECISIONS	
		Offended Violently Group correct separation	Not Offended Violently Group correct separation	High Risk Group correct separation	Low Risk Group correct separation
MODELS USED	Logistic Regression (10 cues)	0%	88%	39%	21%
	Logistic Regression (two most significant cues)	0%	88%	39%	21%
	Matching Heuristic (one cue) with the highest level of accuracy (ie. threat/ control override permitting phenomena)	0%	72%	12%	50%

Using all ten cues in a logistic regression model to predict outcome gave an overall percentage correct figure of 88. This was made up of 88 true negative predictions, and no true positive ones.

The same model used to predict assessors' decisions gave a percentage correct figure of 60. This was made up of 21 low risk and 39 high risk judgements where the model agreed with observed data.

Using only the two most significant cues associated with outcome in a logistic regression model (ie. disorganised social circumstances; use of weapons) gave an overall percentage correct figure of 88. Again, this was made up of 88 true negative predictions, and no true positive ones.

Using only the two most significant cues associated with assessors' decisions in a logistic regression model (ie. history of violence; non-compliance with medication) gave a percentage correct figure of 60. Again, this was made up of 21 low risk and 39 high risk judgements where the model agreed with observed data.

Using the matching heuristic based on only one cue gave levels of accuracy for outcome prediction ranging from 31% to 72%. The best result came from using the threat/control override permitting phenomena cue. This correctly allocated 72% of the study cohort into the not offended violently group, but failed to identify any of the patients who had been found at follow up to have committed violent offences. Using the same cue to predict assessors' decisions gave a percent correct figure of 62. This was made up of 50 people correctly allocated to the low risk group, and 12 correctly placed in the high risk group. The range of variation in accuracy was from 51% to 62%.

Chapter 5

Discussion

5.1 Introduction

It was hoped that this research would show the way to arrive at quick, highly consistent, very specific judgements about which patients were at high risk of committing violence, and who amongst those would go on to offend. Although no models have done this at a better than chance level (see Table Fourteen), it is still worthwhile reflecting on the particular meaning of the associations found.

Of the 100 patients in this study, 46 were assessed as high risk by report writers. Somebody (or some team) had decided that every one of them was provoking enough anxiety and uncertainty to meet criteria for referral on to a specialist (tertiary) service. Yet they were disagreed with by expert opinion in more than half the cases. Of the 54 patients evaluated as low risk, only three went on to offend violently (as false negatives).

The accuracy of predictions that individuals will not be violent when assessed at follow up is known as negative predictive power. The formula for calculating it is;

$$\text{Negative Predictive Power} = \frac{TrueNegatives}{TrueNegatives + FalseNegatives} \times 100$$

(Hart et al, 1993)

For the present study this figure is;

$$\frac{51}{51+3} \times 100 = 94.4\%$$

It appears that Kent Forensic Psychiatry Service assessors were mostly correct in their estimations that patients in the low risk group would not go on to offend violently. Just concluding "low risk" for every patient seen in this sample would yield a success rate in terms of outcome prediction of 88%. A similar conclusion was reached by Steadman (1983) that the likelihood of violent recidivism among a group of offenders is so low that it is difficult to improve on the accuracy of the forecast that nobody will re-offend, even in high risk groups.

The low risk/high risk distinction is however important in terms of resource allocation. It saves case managers from investing staff time in a patient group who are highly unlikely to offend. By contrast, sending a report back to the referral agent that their patient is a high risk of violent behaviour might prompt them and their team to institute measures which could decrease opportunities for the patient to offend. This may have contaminated the study outcome, in that it makes it seem as if the high risk judgement was wrong, and that only 9 out of 46 high risk patients went on to offend (19.56%).

If one considers the proportion of people predicted to be violent, and who actually were so at follow up (i.e. positive predictive power; Hart et al, 1993) the formula is;

$$\text{Positive Predictive Power} = \frac{TruePositives}{FalsePositives + TruePositives} \times 100$$

For the present study this figure is;

$$\frac{9}{37+9} \text{ x } 100 = 19.56\%$$

This is a hit rate for assessors of roughly one in five. This level of success (or sensitivity) does not suggest that the professionals concerned can meaningfully identify the people they most want to find during risk assessment, that is, the true positive cases.

The complement of this figure in standard epidemiological terminology is referred to as the percentage of false positives (Monahan, 1981). It is the proportion of people expected to be violent but who actually were not. The formula is;

$$\text{Percentage False Positives} = \frac{FalsePositives}{FalsePositives + TruePositives} \text{ x } 100$$

For the present study this figure is ;

$$\frac{37}{37+9} \text{ x } 100 = 80.43\%$$

In the five studies Monahan reviewed, the percent false positives ranged from 65% to 86%. For the present study this translates into an error rate of 80.43%. Assessors were wrong in four out of five cases when they judged people to be at high risk of acting violently.

The base rate of violence in the community or cohort under study has a profound effect on prediction figures. In this sample, 12 out of 100 patients had offended violently by time of follow up, while 88 had not. It is therefore more likely that assessors or forecasting models will be able to separate people into the not offended violently group accurately (i.e. specificity) than that they will be able to identify those in the offended violently group (i.e. sensitivity). This holds true for attempts to anticipate any rare event. Geddes (1999) has argued that since homicide and violence are uncommon events, any increase in restrictions placed on those deemed to be at high risk would affect many people who would not go on to behave in an antisocial fashion. Similarly, the report of the National Confidential Inquiry (Appleby et al, 1999) estimates that improving compliance with treatment plans amongst the mentally ill may prevent two homicides per year, while financial and humanitarian costs would be considerable. The impact of a false positive diagnosis cannot be overlooked.

The success rate for true positive predictions by assessors with regard to the offended violently group was nine out of 12, or 75%. However, this figure becomes less impressive when one puts it in the context of the whole study (i.e. nine true positive forecasts compared with 37 patients rated as high risk but who did not go on to offend violently). It compares favourably with true positive predictions from the Gardner et al (1996) study. They developed a screening instrument for use with psychiatric patients involving four variables (i.e. young age, three or more violent acts, heavy illicit drug use, and hostility). This yielded a sensitivity of seven percent (i.e. the proportion of persons predicted to be violent, and who actually were) by

contrast with nine patients out of 46 rated as high risk in the present study (i.e. 19.56%).

The sequence of predictive cues used by postal survey respondents in forecasting violence can be seen in descending order across the second row of Table Three. These are features consultant psychiatrists rated as important when arriving at decisions about the risk a patient poses of behaving violently. Comment will now be made about the frequency of selected cues in the study cohort, correlations amongst them, their relationship to outcomes (i.e. OV or NOV) and decisions (i.e. high or low risk), and how they function in two of the three reasoning models constructed to predict outcome and decision (i.e. logistic regression analysis using two cues; the matching heuristic using only one cue).

5.2 Cue relevance; history of violence

A history of violence is often reported in the research literature as the most reliable and efficient predictor of future violent offending (e.g. McNeil and Binder, 1995; Mossman, 1994; Bonta, Hanson and Law 1998, to cite only a few). Survey respondents in the present study put it at the head of their lists, as may be seen in Table Three, and it was noted in 72% of the sampled cohort. It was observed in 39 of the high-risk patients and 33 of the low risk group. Although present in 11 out of 12 patients who went on after assessment to offend violently, it was also part of the background presentation of 61 people in the group that did not offend violently. It does not therefore make a clinically useful dividing rule for separating these two groups. This is further supported by the fact that using history of violence in a matching heuristic only yielded an accuracy rate of 38% (made up of 11 instances where history of violence and violent offending occurred together, plus 27 instances where there was no history of violence amongst members of the NOV group).

History of violence was correlated with threat/control-override permitting phenomena ($r = 0.203$, $p<0.05$), the use of weapons ($r = 0.260$, $p<0.01$), and risk assessment decision ($r = 0.275$, $p<0.01$). However, it was not significantly associated with violent offending, a result at odds with widely accepted beliefs.

The two-cue logistic regression model based on history of violence and non-compliance with psychotropic medication gave a fit between observed and predicted assessors' decisions of 60%. This was made up of 39% true positive forecasts (ie. high risk judgements), and 21% true negative ones (ie. low risk judgements).

It seems that the data in the present study does not support the belief that a history of violent offending on its own can allow one to make sensitive (i.e. true positive) forecasts about future violence. However, the comparatively small size of the sample in this case limits the strength with which this conclusion can be held.

5.3 Cue relevance; threats or impulses to violence

The cue which occurred most often in the study sample was threats or impulses to violence (81%, as seen in Table Three). This had been ranked third out of 10 predictors by postal survey respondents. One view from the research literature is that threats or impulses to violence have more utility in helping clinicians identify relevant areas of enquiry during risk assessment, rather than serving as indicators of future violence (Grisso et al, 2000). This opinion is supported by data in the present study. The matching heuristic based on threats or impulses to violence gave an accuracy rate in terms of outcome prediction of only 31%. This was made up of 12 instances where threats or impulses to violence and violent offending occurred together, plus 19 cases where there were no such threats or impulses amongst members of the NOV group. This cue was almost equally present in the groups rated high risk (41) and low risk (40) by assessors.

Although present in all 12 of the patients who offended violently, it was also found in 69 of those who did not go on to offend. It is not therefore particularly useful in making sensitive forecasts about violence, but it is no surprise that this feature occurred so frequently in the sampled cohort. Logically one would expect the voicing of threats or observation of impulses to violence as a necessary pre-condition to making a referral on to specialist, tertiary services for an opinion on risk assessment. It is something of a mystery as to what was in the minds of the relevant referral agents when they sent the 19 patients to us who did not demonstrate this feature.

Threats or impulses to violence were correlated with use of weapons ($r = 0.206$, $p < 0.05$), non-compliance with medication ($r = 0.230$, $p < 0.05$), threat/control-overriding permitting phenomena ($r=0.206$, $p < 0.05$), and risk assessment decision ($r = 0.201$, $p < 0.05$). They were also correlated with active symptoms of mental illness ($r = 0.310$, $p < 0.01$), delusions of persecution ($r = 0.387$, $p < 0.01$), availability of victim ($r = 0.322$, $p < 0.01$), and psychiatric aggregate ($r = 0.335$, $p < 0.01$). The latter is an equally weighted aggregate made up of the four predictors which had a psychiatric component (i.e. active symptoms of mental illness, delusions of persecution, non-compliance with a psychotropic medication regime, and threat/control-override permitting phenomena). However, there was no association at a significant level between threats or impulses to violence and violent offending, supporting the opinion that although this cue was present in many cases referred for risk assessment, and that it looks superficially powerful, it should not be given undue weight when making decisions about case management.

5.4 Cue relevance; substance or alcohol misuse history

Postal respondents rated substance or alcohol misuse history second most important amongst the features they looked for when attempting to predict violence, as may be seen in Table Three. In the study sample as a whole it occurred 58 times, spread equally between the high risk (29) and low risk groups (29). It was present in eight out of the 12 violent offenders, but was also reported in 50 of the 88 patients who had not offended violently by follow up. It does not therefore seem to be a useful guide to deciding who will commit a violent offence, a view supported by testing the predictive accuracy of the matching heuristic based on this cue alone. Correct divisions occurred in only 44% of cases, made up of eight instances where substance or alcohol misuse history and violent offending occurred together, plus 36 cases where there was no such history amongst members of the NOV group.

There were no statistically significant correlations between substance or alcohol misuse history and any of the other nine predictors.

These findings are in strong contrast to what has been reported in the research literature. Monahan (2000) wrote that substance abuse was one of the most robust predictors of violence. Blumenthal and Lavender (2000) state that the use of alcohol and drugs often precedes violent offending. Among 26 studies involving 9304 cases, Murdoch, Pihl and Ross (1990) found that 62% of violent offenders were drinking at the time of the offence. Other investigators have found that the majority of violent crimes (e.g. assault, sexual assault, family violence, murder) are committed by individuals who are intoxicated with alcohol (Bonta, Hanson, and Law, 1998; Beck,

1994; Modestin, Burger, and Ammann, 1996). Perhaps sample size and a low base rate of violence in it contributed to this unusual result from this study cohort.

5.5 Cue relevance; availability of a specified victim

This cue was placed ninth out of ten amongst the predictors to violence chosen by consultant psychiatrists in the postal survey, as may be seen in Table Three, but it occurred fourth most frequently in the study cohort (47%). It was found in 22% of the high risk group and 25% of the low risk patients. Half of the 12 people in the offended violently group had specifically identified a potential target individual against whom their aggression might be directed, but so had 41 out of the 88 members who did not go on to commit violent offences after risk assessment.

Using this cue in a matching heuristic yielded an accuracy rate for outcome prediction of 53% (made up of six instances where availability of victim and violent offending occurred together, plus 47 instances where there was no identified potential victim amongst members of the NOV group). This is a better than chance result, and suggests that the ninth ranked cue may be of some limited use in attempting to predict violence.

Availability of a victim was correlated with active symptoms of mental illness ($r = 0.211$, $p<0.05$), psychiatric aggregate ($r = 0.230$, $p<0.05$), and threats or impulses to violence ($r = 0.322$, $p<0.01$). It was not, however, associated in this study cohort with violent offending.

Reviewing the literature shows that the majority of violence by individuals with severe mental disorder is against relatives (most frequently mothers), and that few strangers are attacked (Blumenthal and Lavender, 2000). This finding is supported by Taylor and Gunn (1999), and Bjorkly (1997).

5.6 Cue relevance; use of weapons

This cue was ranked seventh out of ten violence predictors by postal survey respondents, as may be seen in Table Three. It was found in 43 % of the study cohort, divided almost equally between the high risk (21) and low risk groups (22). Eight of the 12 violent offenders had a background of weapons use, but so had 35 out of the 88 people who did not go on to commit violent offences after risk assessment. Using this cue in a matching heuristic yielded an accuracy rate for outcome prediction of 61% (made up of eight instances where use of weapons and violent offending occurred together, plus 53 instances where no use of weapons had been recorded amongst members of the NOV group). This figure is better than one would expect to occur by chance alone, and suggests that the seventh ranked cue may be of some use in attempting to predict violence. This view receives support from Bonta et al (1998), who found that use of, facility with, and collecting of weapons were worthwhile predictors of future violence.

Use of weapons was correlated with threats or impulses to violence ($r = 0.206$, $p<0.05$), and history of violence ($r = 0.260$, $p<0.01$). It was not however associated with violent offending.

The two-cue logistic regression model based on disorganised social circumstances and use of weapons gave an 88% accuracy rate for outcome prediction. This was made up of 88% true negative forecasts, and no true positive ones.

5.7 Cue relevance; disorganised social circumstances

This cue was given last place out of the ten predictors to violence used by consultant psychiatrists in the postal survey, as may be seen in Table Three. It was found in 40% of the study cohort, divided almost equally between the high risk (19) and low risk (21) groups. Eight of the 12 violent offenders were positive for this cue, but so were 32 out of 88 people who did not go on to commit violent offences after risk assessment. Using this predictor in a matching heuristic yielded an accuracy rate for outcome prediction of 60% (made up of eight instances where disorganised social circumstances and violent offending occurred together, plus 52 instances where members of the NOV group were not noted to have disorganised social circumstances). This figure is better than one would expect to occur by chance alone, and suggests that the cue seen as of least importance by survey respondents may actually be of some use when attempting to predict violence.

Support for this view comes from Table Five, where disorganised social circumstances is the only cue to have a significant correlation with violent offending ($r = 0.209$, $p<0.05$).

The two-cue logistic regression model based on disorganised social circumstances and use of weapons gave an 88% accuracy rate for outcome prediction. This was made up of 88% true negative forecasts, but no true positive ones.

In the literature, Gendreau et al (1996) found that family problems and poor living arrangements were related to recidivism. The nature of current intimate relationships was reported by Klassen and O'Connor (1989) as predictive of violence. Using this

factor along with others such as quality of early family life, arrest history, admissions history, and assault as part of the presenting problem allowed the authors to classify correctly 76% of patients into violent or non-violent groups. Social support along with reaction to environmental stressors was viewed by Webster et al (1997b) as an important part of risk estimation. Blumenthal and Lavender (2000) believed that the social context of a person undergoing violence risk assessment, and some appreciation of those they are most likely to interact with, are necessary to make a sound judgement. The authors also claim that the neighbourhood in which an individual resides has been identified as a highly significant predictor of likelihood of violence. Bonta et al (1998) report that a dysfunctional family background is a powerful predictor of violent recidivism. The same authors suggest that because criminal history variables are better forecasters of future offending than are clinical features, risk assessment could be enhanced by paying more attention to the social psychology and criminology literature, while placing less reliance on models of psychopathology.

5.8 Cue relevance; active symptoms of mental illness

The six cues already discussed in detail are related more to personality characteristics than mental illness. It is now appropriate to consider the frequency and salience of the four predictors chosen by postal survey respondents which reflect clinical features and mental state.

Active symptoms of mental illness occurred in 28% of the study cohort, as may be seen in Table Three. The high risk group had 16 patients who were positive for this cue, and the low risk group had 12. Only one of the 12 violent offenders scored on this predictor, while 27 people with this sign appeared in the group which did not go on to offend violently. Using this cue in a matching heuristic yielded an accuracy rate for outcome prediction of 62% (made up of one instance where active symptoms of mental illness and violent offending occurred together, plus 61 occasions where members of the NOV group were not found to have active symptoms). This rate is better than one would expect to occur by chance alone, and suggests that when risk assessors encounter an assessee who does not have active symptoms of mental illness, there is a greater likelihood that they will not go on to offend violently than that they will.

In the literature there is a growing recognition of the limited utility of diagnosis as a risk factor, and an increasing focus on active symptoms of severe mental disorder instead (Blumenthal and Lavender, 2000; Bjorkly, 1997). Monahan (2000) supports this view, and Krakowski et al (1986) came to the conclusion that the frequency of

violent incidents reflects the course of acute psychosis, and is positively correlated with the severity of symptoms. Wessely and Taylor (1991) found that between one and two thirds of homicidal violence committed by individuals with schizophrenia could be attributed to abnormal mental state.

In the present study, active symptoms of mental illness correlated significantly with non-compliance with medication ($r = 0.634$, $p<0.01$), delusions of persecution ($r = 0.576$, $p<0.01$), and threat/control-override permitting phenomena ($r = 0.423$, $p<0.01$). This is no surprise, as one expects to find these features together with florid psychiatric illness.

Active symptoms of mental illness were also associated with threats or impulses to violence ($r = 0.310$, $p<0.01$), and the availability of a specified victim ($r = 0.211$, $p< 0.05$).

5.9 Cue relevance; delusions of persecution

This predictor occurred in 25% of the study cohort, as may be seen in Table Three. In the high risk group 15 patients were positive for this cue, and the low risk group had ten. Only one of the 12 violent offenders scored on this predictor, while 24 with this sign appeared in the group which did not go on to offend violently. Using this cue in a matching heuristic yielded an accuracy rate for outcome prediction of 65% (made up of one instance where delusions of persecution and violent offending occurred together, plus 64 occasions where members of the NOV group were not found to have persecutory delusions). This rate is better than one would expect to occur by chance alone, and suggests that the absence of delusions of persecution would allow a risk assessor to conclude that such a person has a greater likelihood of not going on to offend violently than that they will.

In the literature, Taylor et al (1998) found that 75% of special hospital patients with a diagnosis of functional psychosis were motivated to offend by their delusions.

In the present study, delusions of persecution correlated significantly with non-compliance with medication ($r = 0.449$, $p<0.01$), threat/control override permitting phenomena ($r = 0.338$, $p < 0.01$), and threats or impulses to violence ($r = 0.287$, $p<0.01$).

5.10 Cue relevance; non-compliance with medication

This predictor occurred in 18% of the study cohort, as may be seen in Table Three. The high risk group had twelve patients who were positive for this cue, and the low risk group had six. Only one of the 12 violent offenders scored on this predictor, while 17 people with this sign appeared in the group which did not go on to offend violently. Using this cue in a matching heuristic yielded an accuracy rate for outcome prediction of 71% (made up of one instance where non-compliance with medication and violent offending occurred together, plus 70 occasions where members of the NOV group were not found to have been non-compliant with medication). Although this rate is better than one would expect to occur by chance alone, it does not qualify the cue as a useful discriminator. Little more than a quarter of the sample showed active symptoms of mental illness, and could therefore be expected to have medication prescribed for them, and in the overall United Kingdom population the vast majority of people are not mentally ill. Using this cue on its own would not allow meaningful distinctions to be made between people referred for risk assessment with respect to whether or not they might go on to offend violently. Its sensitivity (i.e. one identified out of 12 violent offenders) would not encourage clinicians to modify their decision making styles.

The two-cue logistic regression model based on history of violence and non-compliance with psychotropic medication gave a fit between observed and predicted assessors' decisions of 60%. This was made up of 39% true positive (or high risk) forecasts, and 21% true negative (or low risk) ones.

In the present study, non-compliance with medication correlated significantly with threat/control-override permitting phenomena ($r = 0.275$, $p<0.01$), and threats or impulses to violence ($r = 0.230$, $p<0.05$).

In the literature, Howlett (1998) estimated that 70% of psychiatric patients discharged from hospital were likely to stop taking their medication within a period of two years. He examined independent enquiry reports into homicide by the mentally ill. Among 35 cases, 20 (or 57%) were found to have non-compliance with medication as a major contributing factor in the breakdown of care prior to the killing.

5.11 Cue relevance; threat/control-override permitting phenomena

The cue which occurred least frequently amongst members of the study cohort was threat/control-override permitting phenomena. It was found in only 15% of cases, as may be seen in Table Three, with 12 of these patients in the high risk group and three in those rated as low risk. None of the 12 violent offenders scored on this predictor. Using it in a matching heuristic yielded an accuracy rate for outcome prediction of 72%, the highest figure for all ten heuristics calculated. This was made up of no instances where threat/control-override permitting phenomena and violent offending occurred together, plus 72 occasions where members of the NOV group were not found to have such phenomena. Although this rate is considerably better than one would expect to occur by chance alone, it does not qualify the cue as a useful discriminator. Noting the absence of a rare sign with zero sensitivity is unlikely to appeal to risk assessors as a meaningful way to predict the likelihood of violence.

In the literature, Swanson et al (1996) found that those who reported threat/control-override symptoms were twice as likely as those reporting other psychotic features to engage in violent behaviour, and about five times as likely as those with no mental disorder. Perhaps this cue has salience only for the particular sub-group of risk assessees who satisfy the criteria for a diagnosis of psychotic illness.

In the present study, threat/control-override permitting phenomena correlated significantly with threats or impulses to violence ($r = 0.206$, $p<0.05$), and history of violence ($r = 0.203$, $p<0.05$).

5.12 Cue relevance: Psyagg

An equally weighted aggregate called Psyagg was made up of the four cues which had a psychiatric component (i.e. active symptoms of mental illness, delusions of persecution, non-compliance with a psychotropic medication regime, and threat/control-override permitting phenomena). This was done to see if such a combination had more power to predict outcomes and decisions than the individual cues themselves.

With regard to outcome (i.e. violent offending), Psyagg could not produce a significant correlation. It was, however, strongly associated with active symptoms of mental illness ($r = 0.864$, $p<0.01$), non-compliance with medication ($r = 0.764$, $p<0.01$), delusions of persecution ($r = 0.772$, $p<0.01$), and threat/control-override permitting phenomena ($r = 0.625$, $p<0.01$). It is no surprise that Psyagg would correlate highly with its own component parts.

It was also associated with risk assessment (i.e. low or high risk categorisation of patients by assessors; $r = 0.344$, $p<0.01$), suggesting that judges in this study see psychiatric signs as important when making decisions about the likelihood of future violence. It is noteworthy that Psyagg was not significantly correlated with violent offending, suggesting that assessors are over-valuing these cues when making predictions.

Risk assessment was significantly correlated with non-compliance with medication ($r = 0.286$, $p<0.01$), threat/control-override permitting phenomena ($r = 0.275$, $p<0.01$),

history of violence ($r = 0.275$, $p<0.01$), and threats or impulses to violence ($r = 0.201$, $p<0.05$). However, none of these predictors were associated with violent offending, again supporting the view that judges in the present study over-emphasised the value of these cues in arriving at their forecasts about who was most likely to offend violently.

5.13 Practical implications for risk assessment

It was hoped that this study would allow the identification of a handful of cues which, being present or absent in a patient's circumstances, would enable categorisation into high or low risk groupings, and specifically discriminate those who go on to offend violently from those who do not. Although this has been possible to some extent, in that one and two cue prediction models have yielded accuracy rates at better than chance levels, sensitivity (i.e. true positive discrimination) is poor. For example, the most successful one cue matching heuristic had an accuracy rate for predicting outcome of 72%, but all of these patients were in the group which had not offended violently at follow up. This is a high level of specificity but shows zero sensitivity. A logistic regression model using two cues produced an 88% accuracy rate, but none of these patients was a violent offender. Using all ten cues in a logistic regression model gave an 88% accuracy rate, but again none of these patients were violent offenders; the people risk assessors most want to identify.

Looking at correlations between predictors and outcome proved just as limited. The only cue associated with violent offending at a statistically significant level was disorganised social circumstances. Unfortunately, this feature was also present in 32 patients who had not offended violently, almost one third of the sample cohort. It does not, therefore, have much clinical utility in arriving at a decision about who is likely to become a violent offender. In the matching heuristic this cue achieved a 60% success rate.

Although of academic interest, these figures for outcome prediction are unlikely to persuade clinicians to change their style of risk assessment.

On a more positive note, assessors in this study classified 54% of the cohort as low risk of violent offending, and only three of these patients went on to prove them wrong. This is a success rate of 94%, and has serious implications for clinical practice. On this basis it is possible to say with a high degree of accuracy that half of the sample are not at particular risk of committing violence, and therefore they do not need further input from professional sources. This allows the scarce clinical resources in place in the community to be directed towards the high risk group, where they are more appropriately needed, and may go some way towards reducing the potential for harm as opposed to the unrealistic idea of attempting to eliminate risk altogether.

Perhaps predicting violent behaviour accurately is impossible for a variety of reasons, such as the low base-rate of violence (12 out of 100 cases in this sample), the interactions between factors which contribute towards the expression of aggression (e.g. beliefs, mental state, environment, other people to name a few), the immediacy of risk, and the potential for a behavioural sequence to be interrupted before it is fully expressed. Should assessors avoid dichotomous decisions such as "this person will/will not offend ", in favour of an approach which identifies factors in the individual and their circumstances which predispose them towards aggression? This would still make risk assessment worthwhile as a way of managing potential for harm, and allow clinicians to work within the boundaries of evidence-based practice. Rather than pretending that we can make accurate predictions with all the

implications they have for the civil liberties of the patient in question, perhaps we should inform referral agents that our role is not to guess what someone might do, but to highlight risk factors which (if left unconsidered) may act as precipitants for violence. This would involve clarifying the mission statements of tertiary services so that those asking for an opinion have a better idea of what to expect. It might also mean including their staff in assessment interviews so they can see that risk assessment proceeds from first principles (not magic), and involves asking the questions others might be afraid to but could if they tried. Arranging regular meetings between key members of referral teams and staff from tertiary services might lead to increased awareness of what each part of the Health Service has to offer, and thereby decrease the frequency of inappropriate requests, as well as educating mental health practitioners in the community about ways they can improve their risk assessment and management strategies.

If people were to claim that they do not have the time for regular meetings to clarify working practice, then perhaps the Kent Forensic Psychiatry Service might mail mission statements, directorate protocols, and patient information check-lists to those who most often use our service. The Tuesday morning allocation meetings regularly become slowed up by staff debating the merits of (and options for responding to) poorly communicated requests. In many cases the data wanted to make decisions is missing or vague, and the authors have to be followed up in person to seek clarification. This wastes time, takes staff away from other tasks, and increases the gap between a patient raising concerns of violent behaviour and being seen for specialist assessment.

Improving the quality and accuracy of communication between those asking for risk assessment and the people providing it is important at a number of levels. The Kent Forensic Psychiatry Service has a county-wide remit covering a population of approximately 1.2 million individuals, which presents resource implications. Funding is limited, the number of secure inpatient beds always seems less than the quantity of requests to put people in them, recruiting and retaining staff is difficult, there is a large geographical area to offer services to, clinical staff report feelings of pressure when attempting to carry out their tasks, and good supervision for practitioners dealing with this challenging patient group is both scarce and costly. Ways around these obstacles can only be found if primary, secondary, and tertiary level healthcare providers make the effort to "talk" with each other, actively listen to opinions expressed, and accept shared responsibility for managing the risks posed by potentially violent patients.

5.14 Limitations of the study

Despite having made every effort to design a methodologically rigorous piece of research, it is inevitable that flaws in it will appear.

This project was not a randomised control trial, where risk assessment requests were put on a waiting list and the results contrasted with those from people actually seen and evaluated. It is difficult to see how such a trial could be conducted ethically, as there is an implicit urgency about every risk assessment referral our service receives.

Patients in the study were chosen as they appeared for assessment, with no distinction made between those who had a diagnosis of mental illness as opposed to those who did not. Although this was an accurate reflection of the population making up the Kent Forensic Psychiatry Service's caseload, perhaps the results might have been different if a selection criterion had been to include only people with a diagnosis of mental illness, or to ensure the cohort had no such people in it.

The predictor variables came from only 37 respondents out of 100 who were written to. Perhaps more refined cues might have come from a much larger survey. The risk assessment reports used as the basis for data collection in this study were written by several different individuals, who had differing levels of experience and theoretical orientation. Consultant Forensic Psychiatrists, Specialist Registrars in forensic psychiatry, and Chartered Clinical and Forensic Psychologists do not necessarily attach the same weights to the 10 predictive cues highlighted. They may

favour some over others, or not even bother to consider items such as disorganised social circumstances when interviewing patients. This variability is a potential limitation of the study which the current design has not been able to quantify. It might have had an impact on the ratio between low and high risk decisions, and we cannot be sure whether judgements were reached on the basis of one, a few, or several predictors having salience for the report writer. This might become clearer if a study was carried out where the reports were provided to a group of assessors who were asked not only to classify patients as high or low risk of committing violence, but also to acknowledge which of the 10 cues they used to reach that decision. There is an ethical consideration to be aware of here. Unlike studies which use fabricated case profiles to identify judgement policies, the risk assessment reports in this project are from interviews with real people. Either the documents would need to be sensitively censored to preserve confidentiality without diluting their content, or such assessors must be chosen from clinical areas and professions where codes of conduct exist to protect patients' rights.

The study does not provide any measure of cross validation. This is a technique for determining the validity of a procedure by testing it for a second time on another sample after its validity has been demonstrated on an initial sample (Chaplin, 1968). The three models used (ie. logistic regression with 10 and two cues, and the matching heuristic) were not used to predict decisions and outcomes for new cases, as this study was a retrospective one, and there was limited time (and only one researcher). With regard to the latter, it was not possible to assess the reliability with which cases were rated (ie. the dependability and consistency with which cues were recognised) because all the data in Appendix B3 was gathered by only one person.

Interviewing the patients might have refined the data collection process and made it more individually accurate, rather than carrying out a retrospective file study. This would involve allowing a considerable gap after information gathering before following up people to see if they had offended violently or not.

The base rate of violence in this sample was low at only 12%, and with any rare event prediction is inherently difficult. One has only to consider the example of television weather forecaster Mr Michael Fish and his 1987 assertion that a hurricane would not arrive in England, to appreciate this. Perhaps increasing the sample size and extending the follow up period would allow more of these unusual events to be captured, and enable us to see if a pattern emerged such that cues individually or in combination were better able to identify true positive cases.

It is possible that the lengthy interviews involved in risk assessment may have played some part in modifying the patient's potential for violence, thus influencing the base rate of offending in the sample. When Hans Eysenk did his pioneering research into the effectiveness of cognitive behaviour therapy in the 1970s, he found that some people who had been interviewed and placed on a waiting list showed symptom remission at follow up despite not having been engaged in treatment.

It may be the case that recommendations contained in the detailed reports by Kent Forensic Psychiatry Service staff might have been acted on by the referral agents, and this could have lowered the frequency of violent behaviour amongst the study cohort. Responsible Medical Officers may have recommended detention under the

appropriate section of the 1983 Mental Health Act. They might have reviewed the impact of medication reported by patients, and changed dosage or type so that clinical benefits exceeded side-effects, thus enhancing compliance with treatment, and increasing the chances of the person's mental state stabilising. Decisions may have been made to offer the person attendance at a day care facility, where particular symptoms could be targeted through occupational therapy, counselling, or specific group work (eg. Anger Management, assertive communication, or social skills training). Community psychiatric nurses might have begun regular home visits, to monitor mental state and level of distress so that potential crises could be identified, defused, and the person's range of options enhanced rather than responding with violence to perceived threat. Social workers may have had their attention drawn to issues such as children at risk in the home, poor financial management, or other practical issues they could offer help with, thereby lowering the level of stress experienced by the person under consideration. If a particular individual had been identified as a possible target for violence, plans might have been made to separate them from the patient, thus changing the likelihood of an offence happening. Reference to current substance misuse could have prompted referral to specialist treatment providers, thereby lowering the risk of disinhibition and clouded judgement. These factors may have accounted (in part) for the low (ie. 12%) rate of violent offending in the study cohort. To measure the impact of requestors' responses to KFPS risk assessment would involve a study design where evaluations were completed, sent to referral agents, sufficient time allowed so there was a reasonable chance of offending occurring, and then further contact made with referrers asking them what steps they had taken to follow the recommendations offered. This could be in the form of a questionnaire with risk minimisation steps

presented, and scored as done or not done. Space would also be provided for general comments. Approaching the Police for guidance about who had offended might then allow us to look for patterns of pre-emptive strategies completed, and their relationship to outcome.

It has been argued that we should abandon altogether the use of 2 x 2 contingency tables when considering the accuracy of violence predictions. Hart et al (1993) claim such matrices are arbitrary and artificial, as neither predictor nor outcome variables divide neatly or naturally into categories, and that reluctance to qualify confidence about judgements is poor practice. Such decisions ought to be presented as based on the balance of probabilities rather than "yes" or "no" forecasts. The authors worry that researchers may be drawn to this type of analysis, and away from multivariate correlational methods which could have greater power to identify links between predictor and outcome variables. Such tables draw the attention of readers to false positives, figures often discouragingly high, which might make investigators reluctant to embark on a quest for new and specific associations which could move the field of violence prediction forward.

Of course, one must never lose sight of the potential for factors other than the ten independent variables considered in this study to have an impact on the expression of violence. The behavioural chain between idea and action is lengthy and complex. Even when one questions violent offenders precisely about an identified incident, it is not easy to isolate causative or precipitating cues. Despite such complexity, and perhaps even because of it, decision makers are still justified in attempting to see if fast and frugal reasoning models (which more plausibly approximate how people

make judgements than do regression ones) can make useful distinctions between those referred for risk assessment about who might go on to offend violently, and who will not.

5.15 Directions for future research

There are a number of ways in which the present study could be built upon and extended.

The format of Appendix B2 could be completed for every risk assessment report done by Kent Forensic Psychiatry Service staff. As things stand at the moment, this would involve drawing colleagues' attention to the wish to collect longitudinal data, and then asking their secretaries to e-mail completed reports to me as soon as they became available. Coding of the ten predictor variables would be done as the document was read, and once sufficient numbers had accumulated, and a reasonable follow up period gone by, contact with the Kent Police could identify who had gone on to offend violently and who had not.

A new cohort could be combined with the existing one to build up a much larger sample size than the current 100, as well as extending the period of time over which people would have had the chance to commit violent offences.

Caution would be necessary if more than one coder was involved. It would be important to ensure that everybody understood the operational definitions of the ten-predictor variables, and was scoring them consistently.

Links could be made with other regional secure units similar to ours and if they had no staff interested in this type of predictive modelling, perhaps agreement might be

obtained for access to their risk assessment reports as they came to hand for coding.

Amalgamating new data with that already obtained would also show if any of the people who had not so far offended changed over to the group who did. If a number did so, it would be interesting to see if the pattern of predictors had changed in some consistent way.

As far as decision making strategies are concerned, the matching heuristic is at the opposite end of the spectrum compared with psychometric instruments such as the Psychopathy Check-List, the Historical/Clinical/Risk Assessment-20, and the Violence Risk Appraisal Guide. Using one or all of these three in parallel with the ten cues already chosen would allow interesting comparisons with regard to sensitivity (i.e. identification of true positive cases), and efficiency (i.e. the amount of time involved in an assessment).

What has become obvious during this study is that referral agents are not always clear about the principal questions they want answered by tertiary services. There may be some benefit in approaching them by way of a survey to quantify how they think the Kent Forensic Psychiatry Service might assist them to identify, manage, and minimise risk of violent behaviour amongst their client group.

At the very beginning of this programme of doctoral study the author considered the dynamic tension between clinical and actuarial approaches to decision making. Some thought was given to reasons for this. It might be helpful to evaluate

practitioner attitudes to moving away from reliance on clinical acumen and the belief that one is expert, towards more empirically based judgement techniques. This could lead to a more healthy degree of humility about what is possible, and give practitioners the evidence they need to justify focusing on risk management strategies, rather than being drawn to the view that if only the right steps are taken then risk can be eliminated.

Chapter 6

Summary And Conclusions

6.1 Purpose of the study

Clinicians in the Kent Forensic Psychiatry Service are often asked by referral agents to make judgements about whether patients living in the community are likely to behave violently or not. There are pressures from high caseloads in a county with approximately 1.2 million inhabitants, limited time to respond to requests, and the usually incomplete nature of the information provided about each patient. Stakes are considerable in terms of the decisions reached. A conclusion of high risk might result in detention for the person concerned against their wishes in a psychiatric hospital under Sections of the 1983 Mental Health Act. Deciding that the person is at low risk of committing violence raises concerns amongst some of those involved in case management that this might be an error, and that somebody may still be victimised. The process currently (i.e. July 2003) in place to make these decisions can take as much as four hours to complete. Case notes must be read, staff and significant others spoken with, the patient themselves interviewed, opinion discussed amongst those in the assessment team, and conclusions formalised into a written report back to the referral agent. All this is based around clinical judgements about the salience of particular cues or risk factors identified during the assessment process, and with rare feedback about whether or not the patient went on to behave violently. Any system which allowed assessors to complete their tasks accurately while avoiding redundant activity could have significant implications for patient care, public safety, resource allocation, and job satisfaction.

These considerations led the author to wonder if a fast and frugal reasoning model could be applied to the prediction of violent offending and decisions about risk level. There was also interest in evaluating the accuracy of this approach against the traditional regression model assumed in so much psychological research to represent the way in which people arrive at judgements by collecting data, weighting it appropriately, and incorporating all the relevant factors to reach an answer.

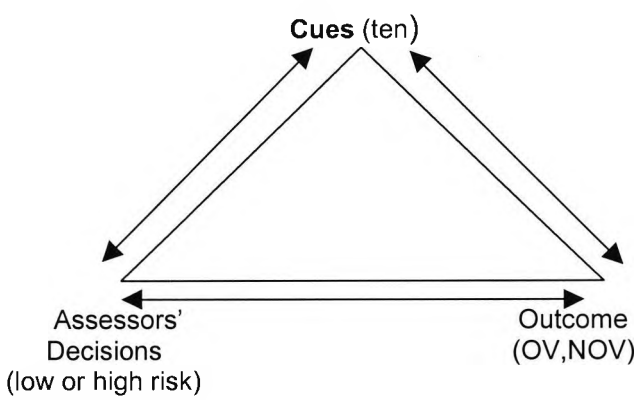
6.2 Summary of methodology

One hundred psychiatrists were contacted through postal survey and asked for the cues they look for when assessing the potential a patient has to commit violence. The answers provided by the 37 respondents were placed in rank order, and the ten most popular (which also received support from the empirical literature on risk prediction) were used as a basis for data collection. One hundred patient risk assessment reports done by Kent Forensic Psychiatry Service staff were checked for the presence or absence of these cues in a retrospective file study. Note was taken of the report writer's risk conclusion at the time (low or high). A search was conducted on the Police National Computer to determine if the patients concerned had subsequently offended violently (OV) or not (NOV).

The following comparisons were possible;

FIGURE 1 INTERACTION BETWEEN STUDY VARIABLES

Adapted from Figure 13.2,
Page 238, in Jackson (1997)



The research hypothesis was that a logistic regression model using ten cues, a logistic regression model using only the two most significant cues, or a matching heuristic using only one cue could accurately separate the 100 patients in this study

into the group which offended violently, the group which did not offend violently, those who were assessed as at high risk of violent offending, and those who were rated as being at low risk of committing violence.

The null hypothesis was that none of the three models used would be able to separate patients accurately into the offended violently, not offended violently, high risk, or low risk groups any better than chance.

6.3 Summary of results

From the study cohort of 100 patients, 46 were assessed as high risk of committing violence, and 54 as low risk. At follow up, 88 had not offended violently, while 12 had done so. One of the offenders was female, and the other 11 were males. Assessors had rated nine of these people as high risk, and three as low risk.

Correlation coefficients were calculated between the ten cues identified as having some relevance when attempting to predict violence. Comparisons were also made between cues, level of assessed risk (i.e. high or low), and with outcome (i.e. offended violently or not). In total 27 correlations achieved statistical significance, eight at the 5% level, and 19 at the 1% level.

The most disappointing feature of this phase of the data analysis from the viewpoint of anyone wanting practical ways to enhance clinical judgement, was the lack of association found between the predictors selected as having some relevance to risk assessment, and the ultimate outcome of violent offending. These ten cues are commonly accepted as evidence on which to make decisions about patient management, and they were readily observable amongst participants in this study, yet they did not have the power to discriminate the OV group from those in the NOV cohort. The most widely cited cue for predicting future violence (namely a history of violence) had a non-significant correlation with violent offending ($r = 0.162$). The other predictors had an apparent relationship to possible violent outcome, but this did not stand up to close scrutiny. None of the correlation coefficients achieved acceptable levels of statistical significance apart from that for disorganised social

circumstances ($r = 0.209$, $p<0.05$). This has implications for the way attempts to predict violence might be conducted in future. Evaluating a patient's social circumstances does not require training in the mental health field. This view was further supported by the lack of association found between violent offending and Psyagg ($r = - 0.179$).

The strong link between Psyagg and risk assessment decisions ($r = 0.344$, $p<0.01$), neither of which correlated with violent offending, suggested that mental health professionals were attending to patient characteristics which had face validity but no bearing on one purpose of the task they were commissioned to do.

The first six cues (which occurred most frequently in this sample) were related more to personality characteristics than mental illness.

The four least frequent predictors were the cues associated with psychiatric conditions.

This suggested that the part played by mental health professionals in conducting risk assessments aimed at the prediction of violence might be overestimated by referring agents. Psychiatric expertise may not add much to the database necessary to reach an accurate conclusion about the likelihood of violent offending.

There was a significant association between prediction of violent offending and outcome ($x^2 = 4.617$; $p=0.032$, $df=1$).

The strength of this link was examined using logistic regression analysis of the study data in an attempt to predict which of the cohort members would fall into the OV or NOV groups. Using all ten cues in a model yielded an 88% correct classification rate, made up of 88 patients who were predicted NOV (and who actually did not offend violently), and no patients who were predicted to offend violently (and who actually did).

A logistic regression model using only the two most significant cues associated with outcome was calculated in order to evaluate its performance against the model using all ten predictors. Utilising disorganised social circumstances and use of weapons revealed a fit between predicted and observed outcome of 88%. There were 88% true negative forecasts and no true positive ones.

Using only one cue in a matching heuristic produced a best level of predictive accuracy for outcomes of 72% (for threat/control-override permitting phenomena). All were true negative forecasts.

Similar approaches were applied to judgements by assessors about the level of risk patients in this study posed.

Using all ten cues in a logistic regression model gave a 60% correct classification rate. This was made up of 21 patients who were predicted as low risk by the model, and who were actually categorised as low risk by assessors (i.e. true negatives). There were 39 patients predicted as high risk and who were also selected as high risk by clinical staff (i.e. true positives).

The two-cue logistic regression model based on history of violence and non-compliance with psychotropic medication gave a fit between predicted and observed decisions of 60%. This was made up of 21 low risk and 39 high-risk judgements.

Using only one cue in a matching heuristic produced a best level of predictive accuracy for decisions about risk of 62% (for threat/control override permitting phenomena). There were 12 true positive forecasts, and 50 true negative ones.

The range of variation for all three models when attempting to predict assessors decisions (i.e. low or high risk) was small at 2%, suggesting that little efficiency was lost by adopting a fast and frugal approach as opposed to a classically rational one. The simpler method was also more plausible as a depiction of how people actually make judgements. For outcome (i.e. OV or NOV group membership) the variation was greater at 16%, but the simpler models still allowed distinctions to be made at a better than chance level.

6.4 Conclusions

An assumption underlying much psychological research about how people make decisions is that this process can be modelled using regression analysis. More recently, the belief that people come to a conclusion after considering all the relevant data, assigning appropriate weights to these factors, and performing a mental calculation to make the optimal choice has been challenged. Fast and frugal reasoning has led to the development of more psychologically plausible probabilistic mental models, which recognise that decisions must be made quickly, and often after incomplete search of the available information. One-choice decision strategies such as the matching heuristic have been found to perform as well as (or better than) classically rational approaches.

The present study set out to compare the accuracy of logistic regression models with the matching heuristic with regard to the prediction of violent offending, and decisions about the level of risk posed by patients referred to the Kent Forensic Psychiatry Service. Postal survey of consultant psychiatrists allowed the identification of ten cues they looked for when attempting to predict violence. One hundred patient risk assessment reports were checked for the presence or absence of these cues in a retrospective file study. Note was made of the report writer's risk conclusion at the time (low or high). A search was conducted on the Police National Computer to determine if the patients concerned had subsequently offended violently or not.

The research hypotheses were that a logistic regression model using ten cues, a logistic regression model using only the two most significant cues, or a matching heuristic using only one cue could accurately separate the 100 patients in this study into the group which offended violently, the group which did not offend violently, those who were assessed as at high risk of violent offending, and those who were rated as being at low risk of committing violence.

The null hypotheses were that none of the three models used would be able to separate patients accurately into the offended violently, not offended violently, high risk, or low risk groups any better than chance.

Table Sixteen shows these combinations, and which of them were supported by the data.

TABLE SIXTEEN: RESEARCH HYPOTHESES AND FINDINGS

		OUTCOME		ASSESSORS' DECISIONS	
		Patients who offended violently	Patients who did not Offend violently	Patients assessed as high risk of committing violence	Patients assessed as low risk of committing violence
MODELS USED	Logistic Regression (10 cues)	ONE accurate separation at a better than chance rate Not supported	TWO accurate separation at a better than chance rate Supported	THREE accurate separation at a better than chance rate Not supported	FOUR accurate separation at a better than chance rate Not supported
	Logistic Regression (two most significant cues)	FIVE accurate separation at a better than chance rate Not supported	SIX accurate separation at a better than chance rate Supported	SEVEN accurate separation at a better than chance rate Not supported	EIGHT accurate separation at a better than chance rate Not supported
	Matching Heuristic using the cue with the highest level of predictive accuracy	NINE accurate separation at a better than chance rate Not supported	TEN accurate separation at a better than chance rate Supported	ELEVEN accurate separation at a better than chance rate Not supported	TWELVE accurate separation at a better than chance rate Not supported

Research hypotheses were supported in terms of predicting membership of the group which did not offend violently. All three models allowed this distinction to be made at a better than chance rate. Levels of accuracy ranged from 72% to 88%. None of the other research hypotheses were supported.

Logistic regression models could only make negative predictions, in other words that all 100 people in the study cohort would not offend violently. None of the independent variables used achieved sufficient levels of statistical significance to allow models to distinguish between those who offended violently (12 people) and those who did not (88 people).

When referral agents ask the Kent Forensic Psychiatry Service for a risk assessment, what they really want is an answer to the question, "Will this individual, with whom I have some statutory obligation to be involved, behave violently, thereby exposing me as a practitioner to accusations that I have failed in my duty of care?" Intuition, comparison between the referred individual and memories of similar people seen previously, structured interview, and actuarial measures all form part of an hierarchical approach to making such a decision. There is however a growing awareness that dichotomous choices are impossible to make with satisfactory levels of accuracy. The results of this study are congruent with others in the research literature. No matter how many cues are considered, either in combination or singly, the relatively rare event of future violence cannot be forecast with adequate sensitivity. Identifying the true positive cases amongst the background clutter from competing signals cannot be done accurately. Even alleged experts rated far more people at high risk of violence than the number who actually went on to offend. It is possible that such a judgement, along with its accompanying feedback to those

managing the patients' care, might have precipitated a series of activities which minimised opportunities for violence to occur. Testing this idea would be a study in its own right. What the reader is left with is the knowledge that from this cohort of patients, all of whom demonstrated features associated with perceived risk of violent behaviour, models from classical and bounded rationality were unable to differentiate true positive cases from all others at a rate which would offer risk assessors useful shorthand tools to make accurate decisions. Perhaps it needs to be made clear to referral agents that probability estimates of violence such as high or low risk will no longer be offered, as there is no sound database to work from. Instead factors specific to the individual under consideration might usefully be highlighted as they relate to potential destabilisation, and a care plan prepared which takes account of their unique contribution to the chances of an undesirable outcome. In this way risk assessment would move away from implied or explicit prediction, and emphasise how care providers might better meet the needs of those in their charge. Evaluations would continue to be time consuming, but more honestly focused on doing what scientist-practitioners are trained for; observing, interpreting, and commenting on people's behaviour rather than pretending their future is predictable.

Section B References

Akers, R.L. (1985). Deviant Behaviour: A Social Learning Approach. Belmont, CA: Wadsworth.

American Psychiatric Association. (1983). Guidelines for legislation on the psychiatric hospitalisation of adults. American Journal Of Psychiatry, Vol. 140, pp 672-679.

Anderson, J.R. (1990). The Adaptive Character Of Thought. Hillsdale, N.J: Erlbaum.

Anderson, C.A. (1989). Temperature and aggression: Ubiquitous effects of heat on occurrence of human violence. Psychological Bulletin, Vol. 106, pp 74-96.

Andrews, D.A. (1982). The Level Of Supervision Inventory (LSI), Toronto, Canada: Ontario Ministry of Correctional Services.

Andrews, D.A., and Bonta, J. (1994). The Psychology Of Criminal Conduct. Cincinnati, Ohio: Anderson.

Appleby, L., Shaw, J., Amos, T., and McDonnell, R. (1999). Safer Services: National Confidential Inquiry Into Suicide And Homicide By People With Mental Illness. London: Department of Health.

Bandura, A. (1983). Psychological mechanisms of aggression, pp 41-69. In R.G. Geen and E.I. Donnerstein (Eds), Aggression: Theoretical And Experimental Reviews, Vol. 1. New York: Academic Press.

Barefoot -v- Estelle, 463 U.S. 880 (1983).

Baxstrom -v- Herald (1966) 383 U.S. 107.

Beck, J.C. (1994). Epidemiology of mental disorder and violence: Beliefs and research findings. Harvard Review Of Psychiatry, Vol. 2, pp 1 – 6.

Bergman, B., and Brismar, B. (1994). Characteristics of violent alcoholics. Alcohol And Alcoholism, Vol. 29, pp 451 – 457.

Berkowitz, L. (1974). Some determinants of impulsive aggression: Role of mediated associations with reinforcements for aggression. Psychological Review, Vol. 81, pp 165-176.

Berkowitz, L. (1993). Aggression: Its Causes, Consequences And Control. New York: McGraw-Hill.

Binder, R., and McNeil, D. (1988). Effects of diagnosis and context of dangerousness. American Journal Of Psychiatry, Vol. 145, pp 728-732.

Bjorkly, S. (1997). Clinical assessment of dangerousness in psychotic patients: Some risk indicators and pitfalls. Aggression And Violent Behaviour, Vol. 2, pp 167 – 178.

Blackburn, R. (1995). Violence. Chapter 4.6, pp 357-373 in Handbook Of Psychology In Legal Contexts, edited by R. Bull and D. Carson, © 1995, John Wiley and Sons Ltd.

Blumenthal, S., and Lavender, T. (2000). Violence And Mental Disorder: A Critical Aid To The Assessment And Management Of Risk. The Zito Trust. Hereford: UK.

Boer, D.P., Hart, S.D., Kropp, P.R., and Webster, C.D. (1998). Manual For The Sexual Violence Risk – 20: Professional Guidelines For Assessing Risk Of Sexual Violence. Vancouver: British Columbia Institute Against Family Violence.

Bonta, J., Hanson, K., and Law (M.), (1998). The prediction of criminal and violent recidivism among mentally disordered offenders: A meta-analysis. Psychological Bulletin, Vol. 123, No. 2, pp 123-142.

Brace, N., Kemp, R., and Snelgar, R. (2000). SPSS for psychologists: A guide to data analysis using SPSS for windows (Versions 8,9, and 10). Hampshire:Palgrave Publishing

Brehmer, B., (1994), The psychology of linear judgement models, Acta Psychologica, Vol. 87, pp 137 – 154.

Breiman, L., Friedman, J.H., Olshen, R.A., and Stone, C.J. (1993). Classification And Regression Trees. New York: Chapman and Hall.

Brooks, A. (1974). Law, Psychiatry, And The Mental Health System. Boston: Little, Brown.

Brunswik, E. (1955). Representative design and probabilistic theory in a functional psychology. Psychological Review, Vol. 62, pp 193 – 217.

Burns, T., and Priebe, S. (1999). Mental health care failure in England: Myth and reality. British Journal Of Psychiatry, Vol. 174, pp 191 – 192.

Chaplin, J.P. (1968). Dictionary of Psychology. Dell Publishing Co., Inc., New York, NY.

Cocozza, J.J., and Steadman, H.J. (1976). The failure of psychiatric predictions of dangerousness: Clear and convincing evidence. Rutgers Law Review, Vol. 29, pp 1084-1101.

Cooke, D.J. (1998). Cross-cultural aspects of psychopathy. In Psychopathy: Theory, Research, And Implications For Society, (Eds. D. Cooke, A.E. Forth, and R.D. Hare). Dordrecht, The Netherlands: Kluwer, pp 13 – 47.

Cook, R.L., and Stewart, T.R. (1975). A comparison of seven methods for obtaining subjective descriptions of judgemental policy. Organizational Behaviour And Human Performance, Vol. 123, pp. 31-45

Cooksey, R.W. (1996). Judgement Analysis : Theory, Methods, and Applications. Academic Press, Inc., San Diego : California

Copas, J., Marshall, P., and Tarling, R. (1996). Predicting reoffending for discretionary conditional release. Home Office Research Study 150. Home Office: London.

Dawes, R.M. (1971). A case study of graduate admissions: Application of three principles of human decision making. American Psychologist, Vol. 26, pp 180 – 188.

Dawes, R.M. (1979). The robust beauty of improper linear models in decision making. American Psychologist, Vol. 34, pp 571 – 582.

Dawes, R.M., Faust, D., and Meehl, P.E. (1989). Clinical versus actuarial judgement. Science, Vol. 243, pp 1668 – 1674.

Department of Health (1990). The Care Programme Approach For People With A Mental Illness Referred To The Specialist Psychiatrist Service. HC (90) 23/LASSL (90) II. London: Author.

Department of Health (1995). Building Bridges. London: HMSO.

Department of Health (1996). The Spectrum Of Care: Local Services For People With Mental Health Problems. London: HMSO.

Developments in the law: Civil commitment of the mentally ill, (1974). Harvard Law Review, Vol. 87, pp 1190-1406.

Dhami, M.K., and Ayton, P. (2001). Bailing and jailing the fast and frugal way. Journal of Behavioural Decision Making. Vol.14, pp. 141-168

Dolan, B., and Coid, J. (1993). Psychopathic And Antisocial Personality Disorder: Treatment And Research Issues. London: Gaskell.

Dolan, M., and Doyle, M. (2000) Violence risk prediction: Clinical and actuarial measures and the role of the Psychopathy Checklist. British Journal Of Psychiatry, Vol. 177, pp 303 – 311.

Dollard, J., Miller, N., Doob, L., Mowrer, O.H., and Sears, R.R. (1939). Frustration And Aggression. New Haven, CT: Yale University Press.

Douglas, K.S., Cox, D.N., and Webster, C.D. (1999). Violence risk assessment: Science and practice. Legal And Criminological Psychology, Vol. 4, pp 149-184.

Eastman, N. (1999). Public health psychiatry or crime prevention? British Medical Journal, Vol. 318, pp 549 – 551.

Farrington, D.P. (1989) Early predictors of adolescent aggression and adult violence. Violence And Victims, Vol. 4, pp 79-100.

Felson, R.B. (1978). Aggression as impression management. Social Psychology, Vol. 41, pp 205-213.

Fuller, J., and Cowan, J. (1999). Risk assessment in a multi-disciplinary forensic setting: Clinical judgement revisited. Journal Of Forensic Psychiatry, Vol. 10, pp 276 – 289.

Gardner, W., Lidz, C.W., Mulvey, E.P., and Shaw, E.C. (1996). A comparison of actuarial methods for identifying repetitively violent patients with mental illnesses. Law And Human Behaviour, Vol. 20, pp 35-48.

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

Gigerenzer, G., Hoffrage, U., and Kleinbolting, H. (1991). Probabilistic mental models: A Brunswikian theory of confidence. Psychological Review, Vol. 98, pp 506 – 528.

Gigerenzer, G. (1993), The bounded rationality of probabilistic mental models, pp 284 – 313, in Manktelow, K.I. and Over, D.E. (Eds.) Rationality: Psychological And Philosophical Perspectives, London: Routledge.

Gigerenzer, G. (1996). The psychology of good judgement: Frequency formats and simple algorithms. Medical Decision Making, Vol. 16, No. 3, pp 273 – 280.

Gigerenzer, G., and Goldstein, D.G. (1996). Reasoning the fast and frugal way: Models of bounded rationality. Psychological Review, Vol. 103, No. 4, pp 650 – 669.

Gilders, I. (1997). Violence in the community: A study of violence and aggression in homelessness and mental health day services. Journal Of Community And Applied Social Psychology, Vol. 7, pp 377 – 387.

Goldberg, L.R. (1965). Diagnosticians versus diagnostic signs: The diagnosis of psychosis versus neurosis from the MMPI. Psychological Monographs, Vol. 79, No. 9, pp 1 – 28.

Gottfredson, D.M., and Bonds, J.A. (1961). A Manual For Intake Base Expectancy Scoring. San Francisco, CA: California Dept. of Corrections, Research Division.

Grann, M. (1998). Personality Disorder And Violent Criminality: A Follow-up Study With Special Reference To Psychopathy And Risk Assessment. Stockholm: Karolinska Institute, Dept. of Clinical Neuroscience and Family Medicine.

Grisso, T., Davis, J., Vesselinov, R., Appelbaum, P.S., and Monahan, J. (2000). Violent thoughts and violent behaviour following hospitalisation for mental disorder. Journal Of Consulting And Clinical Psychology, Vol. 68(3), pp 388 – 398.

Grove, W.M., and Meehl, P.E. (1996). Comparative efficacy of informal (subjective, impressionistic) and formal (mechanical, algorithmic) prediction procedures: The clinical – statistical controversy. Psychology, Public Policy, And Law, Vol. 2, No. 2, pp 293 – 323.

Gunn, J. (1993). Dangerousness. In J. Gunn and P.J. Taylor (Eds.), Forensic Psychiatry: Clinical, Legal, And Ethical Issues. Oxford: Butterworth – Heinemann.

Guttman, L. (1957). Empirical verification of the radex structure of mental abilities and personality traits. Educational And Psychological Measurement, Vol. 17, pp 391-407.

Guttman, L. (1971) Measurement as structural theory. Psychometrika, Vol. 36 (4), pp 329-347.

Hare, R.D. (1991) Manual For The Hare Psychopathy Checklist – Revised. Toronto, Canada: Multi-Health Systems.

Harris, G.T., Rice, M.E., and Quinsey, V.L. (1993). Violent recidivism of mentally disordered offenders: The development of a statistical prediction instrument. Criminal Justice And Behaviour, Vol. 20, pp 315-35.

Hart, S.P., Webster, C.D., Menzies, R.J. (1993). A note on portraying the accuracy of violence predictions. Law And Human Behaviour, Vol. 17, pp. 695-700.

Hart, S.D., Cox, D.N., and Hare, R.D. (1995). The Hare Psychopathy Checklist – Revised Screening Version (PCL:SV). Toronto: Multi-Health Systems.

Hart, S.D. (1998 a). Psychopathy and risk for violence. In Psychopathy: Theory, Research, And Implications For Society. (Eds. D. Cooke, A.E. Forth, and R.D. Hare), pp 355 – 375. Dordrecht, The Netherlands: Kluwer.

Hart, S.D. (1998 b). The role of psychopathy in assessing risk for violence: Conceptual and methodological issues. Legal And Criminological Psychology, Vol. 3, pp 121-137.

Hemphill, J.F., and Hare, R.D. (1996). Psychopathy Checklist factor scores and recidivism. In International Perspectives On Psychopathy, (Eds. D.J. Cooke, A.E. Forth, J Newman, et al), pp 68 – 73. Leicester: British Psychological Society.

Hemphill, J.F., Hare, R.D., and Wong, S. (1998). Psychopathy and recidivism: A review. Legal And Criminological Psychology, Vol. 3, pp 139-170.

Hoffman, P.B. (1983). Screening for risk: A revised salient factor score (SFS:81). Journal Of Criminal Justice, Vol. 1, pp 539 – 547.

Howlett, M. (1998). Medication, Non-compliance, And Mentally Disordered Offenders. London: The Zito Trust.

Jackson, J. (1997) A conceptual model for the study of violence and aggression. In C. Webster and M. Jackson (Eds), Impulsivity: Theory, Assessment, And Treatment, chapter 13, pp 233-247. The Guildford Press: New York.

Kahneman, D., and Tverski, A. (1973). On the psychology of prediction. Psychological Review, Vol 80, pp 237 – 251.

Kahneman, D., Slovic, P., and Tverski, A. (Eds.) (1982). Judgement Under Uncertainty: Heuristics And Biases. Cambridge, England: Cambridge University Press.

Kendall, M.G., 1948a. Rank correlation methods. London:Griffin

Klassen, D., and O'Connor, W.A. (1989). Assessing the risk of violence in released mental patients: A cross-validation study. Psychological Assessment: A Journal Of Consulting And Clinical Psychology, Vol. 1, pp 75-81.

Krakowski, M., Volavka, J., and Brizer, D. (1986). Psychopathology and violence: A review of the literature. Comprehensive Psychiatry, Vol. 27, pp 131 – 148.

Kropp, P.R., Hart, S.D., Webster, C.W., et al (1995). Manual For The Spousal Assault Risk Assessment Guide, 2nd Edition. Vancouver, BC: British Columbia Institute On Family Violence.

Lea, S. (1997). PSY 6003: Logistic regression and discriminant analysis. http://www.exeter.ac.uk/~SEGLea/multrar2/disclogi.html

Levy, S. (1985) Lawful roles of facets in social theories. In D. Canter (Ed.), Facet Theory Approaches To Social Research, pp 59-96. New York: Springer-Verlag.

Lidz, C.W., Mulvey, E.P., and Gardner, W. (1993). The accuracy of predictions of violence to others. Journal Of The American Medical Association, Vol. 269, pp 1007 – 1111.

Link, B.G., and Stueve, A. (1994). Psychotic symptoms and the violent/illegal behaviour of mental patients compared to community controls. In J.Monahan and H.Steadman (Eds), Violence And Mental Disorder: Developments In Risk Assessment, Chicago, Illinois, USA: Chicago University Press.

Link, B.G., and Stueve, A. (1995). Evidence bearing on mental illness as a possible cause of violent behaviour. Epidemiologic Reviews, Vol. 17, pp 172 – 181.

Litwack, T.R. (1994). Assessments of dangerousness: Legal, research and clinical developments. Administration And Policy In Mental Health, Vol. 21, pp 361 – 377.

Lorenz, K. (1966). On Aggression. London: Methuen.

Mark, V.H., and Ervin, F.R. (1970). Violence And The Brain. New York: Harper and Row.

Marshall, P. (1997). A Reconviction Study Of HMP Grendon Therapeutic Community. London: HMSO.

Marzuk, P.M. (1996). Violence, crime and mental illness. How strong a link? Archives Of General Psychiatry, Vol. 49, pp 472 – 475.

McClosky, D.N., (1985), The Rhetoric Of Economics. Madison: University of Wisconsin Press.

McNiel, D.E., Binder, R.L., and Greenfield, T.K. (1988). Predictors of violence in civilly committed acute psychiatric patients. American Journal Of Psychiatry, Vol. 145, pp 965-970.

McNiel, D.E., and Binder, R.L. (1991). Clinical assessment of the risk of violence among psychiatric inpatients. American Journal Of Psychiatry, Vol. 148, pp 1317-1321.

McNiel, D.E., and Binder, R.L. (1995). Correlates of accuracy in the assessment of psychiatric inpatients' risk of violence. American Journal Of Psychiatry, Vol. 152, pp 901-906.

Meehl, P.E. (1954). Clinical Versus Statistical Prediction, A Theoretical Analysis And A Review Of The Evidence. University of Minnesota Press, Minneapolis, MN.

Michon, J.A., and Pakes, F.J. (1995). Judicial Decision-Making: A Theoretical Perspective, Chapter 6.2, pp 509 – 525, in Bull, R. and Carson, D. (Eds.) Handbook Of Psychology In Legal Contexts, 1995, John Wiley and Sons Ltd.

Modestin, J., Berger, A., and Ammann, R. (1996). Mental disorder and criminality: Male alcoholism. Journal Of Nervous And Mental Disorder, Vol. 184, pp 393 – 402.

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

Monahan, J. (1984). The prediction of violent behaviour: Toward a second generation of theory and policy. American Journal Of Psychiatry, Vol. 141, pp 10-15.

Monahan, J. (1992). Mental disorder and violent behaviour: Perceptions and evidence. American Psychologist, Vol. 47, pp 511 – 521.

Monahan, J. (1993 a). Causes of violence. In U.S.S. Commission (Ed.), Drugs And Violence In America, Washington, D.C.: U.S. Government Printing Office.

Monahan, J. (1993 b). Limiting therapist exposure to Tarasoff liability: Guidelines for risk containment. American Psychologist, Vol. 48, pp 242 – 250.

Monahan, J., and Steadman, H.J. (Eds) (1994). Violence And Mental Disorder: Developments In Risk Assessment. Chicago, Illinois: University of Chicago Press.

Monahan, J. (1996). Violence prediction: The past twenty and the next twenty years. Criminal Justice And Behaviour, Vol. 23, pp 107-120.

Monahan, J., and Steadman, H.J. (1996). Violent storms and violent people: How meteorology can inform risk communication in mental health law. American Psychologist, Vol. 51, pp 931-938.

Monahan, J., Steadman, H.J., Appelbaum, P.S., Robbins, P.C., Mulvey, E.P., Silver, E., Roth, L.H., and Grisso, T. (2000). Developing a clinically useful actuarial tool for assessing violence risk. British Journal of Psychiatry, Vol. 176, pp 312-319.

Moore, B. (1996). Risk Assessment: A Practitioner's Guide To Predicting Harmful Behaviour. London, UK: Whiting and Birch Ltd.

Morrison, E.F. (1994) The evolution of a concept: Aggression and violence in psychiatric settings. Archives Of Psychiatric Nursing, Vol. 8, pp 245 - 253.

Mossman, D., and Somoza, E. (1992). Balancing risks and benefits: Another approach to optimizing diagnostic tests. Journal Of Neuropsychiatry And Clinical Neuroscience, Vol. 4, pp 331 – 335.

Mossman, D. (1994). Assessing predictions of violence: Being accurate about accuracy. Journal Of Consulting And Clinical Psychology, Vol. 62, pp 783-792.

Moyer, K. (1981). Biological substrates of aggression: Implications for control, pp 15-32. In P.F. Brain and D. Benton (Eds), The Biology Of Aggression. Alphen aan der Rijn: Sijthoff and Noordhoff.

Mulvey, E.P., and Lidz, C.W. (1985). Back to basics: A critical analysis of dangerousness research in a new legal environment. Law And Human Behaviour, Vol. 9, pp 209 – 219.

Mulvey, E.P., and Lidz, C.W. (1995). Conditional prediction: A model for research and dangerousness to others in a new era. International Journal Of Law And Psychiatry, Vol. 18, pp 129 – 143.

Murdoch, D., Pihl, R.O., and Ross, D. (1990). Alcohol and crimes of violence: Present issues. International Journal Of The Addictions, Vol. 25, pp 1065 – 1081.

National Centre for State Courts. (1986). Guidelines for involuntary commitment. Mental And Physical Disability Law Reporter, Vol. 10, pp 409-514.

National Research Council (1989). Improving Risk Communication. Washington, DC: National Research Press.

Nuffield, J. (1989). The "SIR Scale": Some reflections on its applications. Forum On Corrections Research, Vol. 1, pp 19-22.

Olweas, D. (1979). Stability of aggressive reaction patterns in males: A review, Psychological Bulletin, Vol. 86, pp 852-875.

Onyett, S.R., and Ford, R. (1996). Community mental health teams – where's the wreckage? Journal Of Mental Health, Vol. 5, pp 47 – 55.

Onyett, S.R. (1999). Community mental health team working on a socially valued enterprise. Journal Of Mental Health, Vol. 8, No. 3, pp 245 – 251.

Performance Management Directorate (1993). Health Of The Nation: Mentally Disordered Offenders. Leeds: NHS Management Executive.

Pihl, R.O., and Peterson, J.B. (1993). Alcohol/drug use and aggressive behaviour. In S. Hodgins (Ed), Mental Disorder And Crime, Newbury Park, CA, USA: Sage Publications.

Quinsey, V.L., Harris, G.T., Rice, G.T., and Cormier, C.A. (1998). Violent Offenders: Appraising And Managing Risk. Washington DC: American Psychological Association.

Reber, A. (1985). Dictionary Of Psychology. Harmondsworth, UK: Penguin Books.

Reith, M. (1998). Risk assessment and management: Lessons from mental health inquiry reports. Medicine, Science, And The Law, Vol. 38, pp 221 – 226.

Rice, M.E., Harris, G.T., and Cormier, C. (1992). Evaluation of a maximum security therapeutic community for psychopaths and other mentally disordered offenders. Law And Human Behaviour, Vol. 16, pp 399 – 412.

Rice, M.E., and Harris, G.T. (1995). Violent recidivism: Assessing predictive validity. Journal Of Consulting And Clinical Psychology, Vol. 53, pp 737 – 748.

Rice, M.E., and Harris, G.T. (1997). Cross validation and extension of the Violence Risk Appraisal Guide for child molesters and rapists. Law And Human Behaviour, Vol. 21, pp 231 – 241.

Royal College of Psychiatrists, Membership List (2000), Bell and Bain Ltd, Glasgow.

Salekin, R.T., Rogers, R., and Sewell, K.W. (1996). A review and meta-analysis of the Psychopathy Checklist and Psychopathy Checklist – Revised: Predictive validity of dangerousness. Clinical Psychology: Science And Practice, Vol. 3, pp 203-215.

Schall -v- Martin, 104 S. Ct. 2403 (1984).

Shannon, C.E., and Weaver, W. (1949). Mathematical Theory Of Communication. Urbana: University of Illinois Press.

Shaw, J., Appleby, L., Amos, T., McDonnell, R., Harris, C., McCann, K., Kiernan, K., Davies, S., Bickley, H., and Parsons, R. (1999). Mental disorder and clinical care in people convicted of homicide: National clinical survey. British Medical Journal, Vol. 318, pp 1240 – 1244.

Sheppard, D. (1996) Learning The Lessons. London: Zito Trust.

Shore, D., Filson, C.R., Johnson, W.E., Rae, D.S., Mueher, P., Kelley, D.J., Davis, T.S., Waldman, I.N., and Wyatt, R.J. (1989). Murder and assault arrests of White House cases: Clinical and demographic correlates of violence subsequent to civil commitment. American Journal Of Psychiatry, Vol. 146, pp 645 – 651.

Siegel, S. (1956). Nonparametric Statistics For The Behavioural Sciences. McGraw-Hill, Kogakusha, Ltd.

Simon, H.A. (1955). A behavioural model of rational choice, Quarterly Journal Of Economics, Vol. 69, pp 99 – 118.

Simon, H.A. (1956). Rational choice and the structure of the environment. Psychological Review, Vol. 63, pp 129 – 138.

Simon, H.A. (1982). Models Of Bounded Rationality. Cambridge, M.A: MIT Press.

Simon, H.A. (1990). Invariants of human behaviour. Annual Review Of Psychology, Vol. 41, pp 1 – 19.

Slaby, R.G., and Guerra, N.G. (1988). Cognitive mediators of aggression in adolescent offenders: 1 Assessment. Developmental Psychology, Vol. 24, pp 580-8.

Snowden, P. (1997). Practical aspects of clinical risk assessment and management. British Journal Of Psychiatry, Vol. 170 (Supplement 32), pp 32 – 34.

Steadman, H.J., and Cocozza, J.J. (1974). Careers Of The Criminally Insane: Excessive Social Control Of Deviance. Lexington, MA: Lexington Books.

Steadman, H.J. (1983). Predicting dangerousness among the mentally ill: Art, magic and science. International Journal Of Law And Psychiatry, Vol. 6, pp 381-390.

Steadman, H.J., Mulvey, E., Monahan, J., Robbins, P.C., Appelbaum, P.S., Grisso, T., Roth, L.H., and Silver, E. (1998). Violence by people discharged from acute psychiatric inpatient facilities and by others in the same neighbourhoods. Archives Of General Psychiatry, Vol. 55, pp 393-401.

Steadman, H.J., Silver, E., Monahan, J., Appelbaum, P.S., Robbins, P.C., Mulvey, E.P., Grisso, T., Roth, L.H., and Banks, S. (2000) A classification tree approach to the development of actuarial violence risk assessment tools. Law And Human Behaviour, Vol. 24, No. 1, pp 83-100.

Swanson, J.W. (1993). Alcohol abuse, mental disorder, and violent behaviour: An epidemiological inquiry. Alcohol, Health And Research World, Vol. 17, pp 123 – 132.

Swanson, J.W., Borum, R., Swartz, M.S., and Monahan, J. (1996). Psychotic symptoms and disorders and the risk of violent behaviour in the community. Criminal Behaviour And Mental Health, Vol. 6, pp 309 – 329.

Tarasoff -v- Regents of the University of California, 551 P. 2d 334 (1976).

Taylor, P. (1982). Schizophrenia and violence. In Abnormal Offenders, Delinquency, And The Criminal Justice System. J. Gunn and D.P. Farrington (Eds), pp 269-284. New York: Wiley.

Taylor, P.J., and Gunn, J. (1984). Violence and psychosis: Risk of violence among psychotic men. British Medical Journal, Vol. 288, pp 1945 – 1949.

Taylor, P.J. (1985). Motives for offending among violent and psychotic men. British Journal Of Psychiatry, Vol. 147, pp 491 – 498.

Taylor, P.J., Garety, P., Buchanan, A., Reed, A., Wessely, S., Ray, K., Dunn, G., and Grubin, D. (1994). Delusion and violence. In J. Monahan and H. Steadman (Eds), Violence And Mental Disorder: Developments In Risk Assessment, Chicago, Illinois, USA: Chicago University Press.

Taylor, P.J., Leese, M., Williams, D., Butwell, M., Daly, R., and Larkin, E. (1998). Mental disorder and violence: A special (high security) hospital study. British Journal Of Psychiatry, Vol. 172, pp 218 – 226.

Taylor, P.J., and Gunn, J. (1999). Homicides by people with mental illness: Myth or reality. British Journal Of Psychiatry, Vol. 174, pp 9 – 14.

Thornberry, T.P., and Jacoby, J.E. (1979). The Criminally Insane: A Community Follow-up Of Mentally Ill Offenders. Chicago, Illinois: University of Chicago Press.

Timmins, N. (1996). The Five Giants: A Biography Of The Welfare State. London: Harper Collins.

Toch, H. (1969). Violent Men. Harmondsworth: Penguin.

Torrey, E.F. (1994). Violent behaviour by individuals with serious mental illness. Hospital And Community Psychiatry, Vol. 45, pp 653 – 662.

Valenstein, E.S. (1976) Brain stimulation and the origin of violent behaviour, pp 101-24. In W.L. Smith and A. Kling (Eds) Issues in Brain/Behaviour Control. New York: Spectrum.

Van Koppen, P.J. (1995). Judges' Decision-Making, chapter 6.7, pp 581 – 610, in Bull, R. and Carson, D. (Eds.), Handbook Of Psychology In Legal Contexts, 1995, John Wiley and Sons Ltd.

Wagenaar, W.A., Van Koppen, P.J., and Crombag, H.F.M. (1993). Anchored Narratives: The Psychology Of Criminal Evidence. Hemel Hempstead, UK: Harvester Wheatsheaf.

Webster, C.D., Harris, G.T., Rice, M.E., et al (1994). The Violence Prediction Scheme: Assessing Dangerousness In High Risk Men. Toronto: University of Toronto, Centre of Criminology.

Webster, C.D., Menzies, R.J., and Hart, S.D. (1995). Dangerousness and risk. In R. Bull and D. Carson (Eds.), Handbook Of Psychology In Legal Contexts, pp 465 – 479. Chichester: John Wiley and Sons Ltd.

Webster, C.D., Douglas, K.S., Eaves, D., et al (1997 a). Assessing risk of violence to others. In Impulsivity: Theory, Assessment, And Treatment. (Eds. C.D. Webster and M.A. Jackson). New York: Guilford Press.

Webster, C.D., Douglas, K.S., Eaves, D., and Hart, S.D. (1997 b). HCR-20: Assessing Risk For Violence (Version 2). Vancouver: Mental Health, Law and Policy Institute, Simon Fraser University.

Wessely, S., and Taylor, P.J. (1991). Madness and crime: Criminology versus psychiatry. Criminal Behaviour And Mental Health, Vol. 1, pp 193 – 228.

Wessely, S. (1993). Violence and psychosis. In C. Thompson and P. Cowen (Eds.), Violence: Basic And Clinical Science, (pp 119 – 134). Oxford, UK: Butterworth – Heinemann.

Wilkening, H.E. (1973). The Psychology Almanac: A Handbook For Students. Brooks/Cole Publishing Company, Monterey, California.

Wilson, D., and Ashton, J. (1998). Crime And Punishment. London: Blackstone Press Ltd.

Wolfgang, M.E. (1958) Patterns In Criminal Homicide, Philadelphia: University of Pennsylvania Press.

Wolfgang, M.E., and Ferracutti, F. (1967). The Subculture Of Violence. London: Tavistock.

Zamble, E., and Palmer, W. (1996). Prediction of recidivism using psychopathy and other psychologically meaningful variables. In International Perspectives On Psychopathy, (Eds. D.J. Cooke, A.E. Forth, J Newman, et al), pp 153-156. Leicester: British Psychological Society.

Zillman, D. (1978). Hostility And Aggression. Hillsdale, New Jersey: Erlbaum Associates.

Section C: Case study

Chapter 7

When preoccupation intrudes on personal freedom: A report on work with a media personality stalker

7.1 Introduction

This report describes the assessment, treatment, management, and follow up of an individual (alias Mr Field) convicted under the anti-stalking laws and sentenced to imprisonment. Consideration will be given to the principle theoretical orientation of the psychotherapist (Rational Emotive Behavioural Counselling), how that relates to the client and his difficulties, and how this particular type of intervention differs from commonly accepted psychological practice. The Rogerian Concept of client-centred therapy (Rogers, 1951) and its core belief that practitioners must demonstrate unconditional positive regard (Truax, 1963, 1966) have shaped therapeutic interventions in a humanistic way, and are key components in many diverse treatment modalities, but when dealing with a forensic client group one must also consider how personal freedoms interact with the rights of others. Balancing the needs of client and community is challenging, and highlights the importance of peer support, consultation, and supervision.

7.2 Summary of theoretical orientation

There are as many ways of approaching psychological problems as there are problems themselves, and the practitioner must therefore make choices about how to apply what they know to the challenge of therapy. Early in my post graduate clinical training (1977) I was introduced to the concepts of Rational –Emotive Therapy (RET) as put forward by Albert Ellis (Ellis, 1962). He was not, of course, the first to discover this way of construing the world and people's ways of reacting to what happens in it. The stoic philosopher Epictetus observed in the first century AD that "people are disturbed not so much by events as by the views which they take of them" (Scott & Dryden, 1996, Page 156). This theme was echoed by William Shakespeare when he wrote, "there is nothing either good or bad but thinking makes it so".

Ellis' approach is now known as Rational Emotive Behavioural Counselling (REBC, Nelson-Jones, 1995a), and strongly emphasises the interplay of feeling, behaviour, and cognition as people try to achieve the fundamental goals of surviving in a hostile world while simultaneously avoiding pain and being happy. There are biological tendencies which push us towards the successful accomplishment of our plans as well as predisposing us on occasions to think and behave irrationally. There is, however, always some element of free choice.

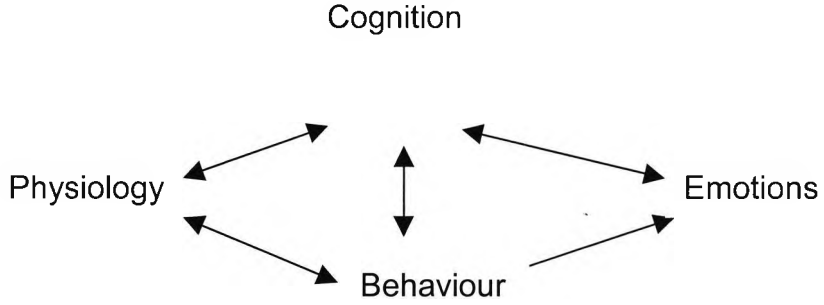
Irrational beliefs are innate, modified partly by social learning, but also come about because people do not develop and exercise their capacity for rational choice. The presence of such ideas can lead the individual to develop an idealised view of the

world and those around them which is seldom realised. The gap between what one expects and what actually happens leads to disappointment, and when this process is repeated continually psychological distress can develop. A scientist who has shed some light on how this might come about is Hans Selye. In his 1956 book (The Stress of Life) he outlined the general adaptation syndrome. The organism first reacts to a stressor with alarm but then moves into a phase of resistance to overcome whatever the problem is. If this does not lead to a satisfactory resolution other options are tested, but if they too fail then exhaustion sets in. People find it difficult to persevere in the face of adversity, and often give up, leaving themselves with the belief that what they are feeling is directly linked to the challenging event or stressor which they have just grappled with. REBC proponents see this as a mistake, and explain to their clients that in between the Activating event and the Consequences (both emotional and behavioural) lie a variety of Beliefs which can be rational or irrational (Nelson-Jones, 1995a). Rational beliefs are associated with adaptive coping while irrational ones lead to problems. Rather than put up with this and view one's fate as outside of the individual's control, REBC teaches people how to Dispute irrational beliefs once they are identified, and to incorporate their results into an Effective new philosophy. Scientific questioning is the main cognitive disputing technique, but use is also made of behavioural assignments as ways for the client to practice coping strategies between therapeutic sessions, and to highlight areas of misunderstanding which might warrant further input. The principal goals of REBC are to help people overcome emotional blocks and disturbances so they can function better. The counsellor is a teacher using a variety of cognitive, emotive, and behavioural techniques to help clients Dispute irrational beliefs (and their derivatives) so they can develop Effective new philosophies to achieve their Goals in life.

The framework therapists use when working with clients is to examine the sequence and links between Goals, Activating events, Beliefs, Consequences, Disputation of irrational beliefs, and the development of Effective new philosophies to live by (Nelson-Jones, 1995a). A shorthand mnemonic for this process is GABCDE.

The interactive nature of REBC means that it does not stand alone, and is seen by many as a specialised part of the larger field of cognitive-behaviour therapy (Scott & Dryden, 1996). The theoretical origins of this approach stem from the view Epictetus took of the world and how we fit in it. This may be illustrated as follows:

Figure 4 RATIONAL EMOTIVE BEHAVIOURAL COUNSELLING FRAMEWORK



Such a framework gives both therapist and client a variety of starting points from which they can challenge ineffective coping strategies and thinking styles.

7.3 The context for the work

I am employed by the Kent Forensic Psychiatry Service (KFPS) based at the Trevor Gibbens Unit (TGU) in Maidstone. This is a medium secure unit with a total of 32 inpatient beds. It is a part of the British National Health Service (NHS) rather than the prison service. The unit philosophy is primarily care based, with security as an important (but secondary) consideration.

In 1996 our service directorate was successful in bidding for a contract to offer mental health consultation to the prisons in the Kent County Cluster (Wilkinson, 1998). On a regular (usually weekly basis) psychiatrists, psychologists, and community psychiatric nurses hold clinics on site for inmates. I go each week to HMP Swaleside and HMP Elmley on the Isle of Sheppey, and to HMP Rochester on the mainland. Clinic space is at a premium and often results in appointments being as much as three or four weeks apart. The aim is for a 50 minute session with each person, allowing ten minutes for note writing before seeing the next one. It is usual to plan for three people to be seen at each visit. The prison service treats outside contractors with some respect as far as practical matters involved in running clinics are concerned, but reserves the right to use or ignore our advice as healthcare governors and their managers see fit. There is no guarantee that a prisoner will be in a particular institution for as long as it takes to complete treatment, and this puts pressure on the therapeutic relationship. It is important to plan ahead, to keep paper work scrupulously up to date, and to liase closely with prison, probation, and health care staff. Each session needs to be conducted as if it might be the final one,

although sometimes, as in this case study, it is possible to maintain contact with the inmate/client from one prison to another, and eventually back into the community.

7.4 The referral

Mr Field was initially assessed by an independent Consultant Forensic Psychiatrist at the request of his barrister. One of the findings was that it would be helpful to commission a further evaluation, this time by a Consultant Forensic Psychologist, as it seemed there were issues of a psychotherapeutic nature to be addressed rather than mental illness. Reports from both these people were on file at the prison by the time Mr Field was sentenced to 30 months' imprisonment for making threats to kill, harassment with threats, and causing fear of violence. This group of charges has come to be known under the generic description of stalking. There is no consensus about the exact definition of stalking (Kamphuis & Emmelkamp, 2000), with most disagreement focusing on the degree of emphasis placed on the extent to which the behaviour evokes a subjective sense of threat. Meloy & Gothard (1995) offer a useful way to view this type of conduct; "Stalking is typically defined as the wilful, malicious, and repeated following or harassing of another person that threatens his or her safety", (Page 258).

I first met Mr Field on Thursday January 20th 2000 as part of a joint interview with a locum Consultant Forensic Psychiatrist. This session had been arranged in response to a request from healthcare staff at the prison who were aware of the professional opinions already on file. The psychiatrist led the assessment which lasted around two hours. By its conclusion we were satisfied that there was no mental illness which needed psychiatric intervention and oversight, but that Mr Field spoke in such a way as to convince us that he held a number of irrational beliefs which had contributed significantly to his offending behaviour. Using his mental

state, personal history, educational achievements, and opinions about how he had come to be in prison, the psychiatrist and I agreed that Mr Field needed (and had the potential to benefit from) psychological treatment. REBC was seen as a logical approach in this instance, and I agreed to meet with Mr Field at the next available clinic time on Friday February 25th 2000. This was the first of sixteen individual therapy sessions, ending on Monday April 23rd 2001.

7.5 The presenting problem: Initial assessment and formulation

Mr Field held beliefs about himself, relationships, and need gratification which led him to act outside the law, resulting in imprisonment. With no friends and nobody to give him feed-back on appropriate social interactions, his fantasy life was able to expand beyond normal social boundaries, and he pursued his preoccupation with the victim past the limits which reasonable people would find acceptable.

The beliefs which Mr Field held about the victim and how to interact with her were not of a delusional nature. Solomon & Patch (1974) define this type of thinking as involving "a false belief, usually unique to the individual, which is not susceptible to modification or correction by logical persuasion or by compelling contrary evidence",(p.68). His interest in the victim was more addictive in nature than classically obsessive and compulsive. He pursued her despite knowing that it was harmful to him in that he neglected his university studies as a result, he craved vicarious contact with her via the post or watching her on television, and he missed this contact if it was thwarted for any reason (ICD-10, 1993). These three manifestations occurred together for at least one month, and satisfy the criteria for a diagnosis of dependence syndrome (p.57).

At first glance, Mr Field might seem to be suffering from an obsessive compulsive disorder. In psychiatric terms, an obsession is a thought or belief which is unwelcome to the person and resisted by them but which keeps intruding on their consciousness (Ullman & Krasner, 1969). A compulsion is the physical

manifestation of an obsession. While the individual knows such activities are nonsensical, they are compelled to do them in order to avoid being overwhelmed by anxiety. None of these features applied to Mr Field. His preoccupation with the victim was welcome, satisfying, enjoyed, and he rarely tried to resist it. He did not see it as inappropriate but felt it was a reasonable way to proceed.

In terms of his personality, Mr Field described (and his parents confirmed) longstanding traits of a schizoid, narcissistic, and obsessional nature. The schizoid traits were evidenced by his lack of interest in close and confiding relationships, his social withdrawal and isolation, and his tendency to see people as irrelevant or as objects. This was shown in part by his preference for a fantasy relationship with the victim rather than a real one with an actual partner.

Narcissistic traits were evident in his beliefs about his sexual prowess, his intellectual abilities, and his difficulty seeing things from other people's points of view. His world was centred around himself. He believed that as he knew in his own heart that he did not intend the victim any harm, then it was unreasonable for anybody else to infer otherwise. Mr Field could not see his actions from the perspective of the victim or a disinterested observer.

Obsessional traits were evident in a rigidity of behaviour and an insistence on things being done a particular way. Mr Field has approached a wide variety of interests over the years as "projects". He has become excessively preoccupied with subjects,

studied and collected material about them at the expense of other commitments until becoming bored, and then moved on to something else. He reported a need to "be obsessed" with something in order to feel that his life had purpose.

The above factors have been reviewed because when someone presents in as unusual a fashion as Mr Field it is important to establish the presence or absence of a significant mental illness. This has obvious implications for the direction intervention might take. When one considers how Mr Field fared at interview, the absence of a previous psychiatric illness history, and reports from the family that his mental state had not changed significantly at any stage, then it is reasonable to conclude that there is no active mental illness. However, there was ample evidence that Mr Field's personality included schizoid, narcissistic, and obsessional traits. Their combination was sufficient to have a direct relationship on his actions towards the victim. He began to lead an increasingly isolated existence, detached from the objective checks of contact with others and the impact this might have on his assessment of reality. The only rules which applied were of his own making. The focus on the victim grew out of his general interest in female television presenters. His life at home when fascination with the victim began consisted mainly of watching television and working on his computer. There had been no significant social contacts outside the home for over two years. Mr Field rarely emerged from his room other than to collect his meals. Under these circumstances his only potential sexual or emotional partners were people he watched on television. Initially, his interest in the victim was slightly inappropriate but not sufficiently so as to warrant concern. As with many previous projects, his fascination with the victim grew to

exclude other activities, and became an abnormal preoccupation. Lacking the boundaries of normal social interactions and receiving no feed back on his behaviour, Mr Field pursued this involvement to unacceptable lengths. He became intellectually, emotionally, and sexually obsessed with the victim.

His personality characteristics did not meet the criteria for a diagnosis of psychopathy under the terms of the Mental Health Act (1983), as they had not led Mr Field to behave in a seriously irresponsible or abnormally aggressive way in any context other than his preoccupation with the victim. In other aspects of his life his functioning remained largely intact. This conclusion, along with the absence of mental illness, meant that it would not be appropriate for Mr Field to be admitted to a psychiatric facility under any section of the Mental Health Act (1983). Although he was seen by the court as clearly in need of containment and psychological therapy, the consensus was that he was not suffering from sufficient impairment to his judgement from mental health problems that he needed to be placed in hospital for his own good or that of others.

7.6 Biographical details

At the time of writing (i.e. 26/04/01) Mr Field was 31 years old. He was born and raised in the South of England, and has one sibling (an older sister). She is single with no children and works in business management in London. The parents are white, healthy, and, from Mr Field's account, comfortably married. The father is retired from retail work and restores antique clocks as a hobby. Mr Field's mother works for a local government agency. None of the immediate family have past or present mental or physical health concerns. There was a maternal uncle who was seen as eccentric, and who died of natural causes in 1990. The family have commented that Mr Field is very similar to this man.

A psychiatric report prepared for the court states that Mr Field experienced an unremarkable birth and early development. The family have noted that he reached the milestones of infancy (i.e., walking, talking, toilet training) ahead of his peers. He was described as a well-behaved child. At primary school he received glowing reports from his teachers, and presented as a quiet, studious boy. It was a surprise to those who knew him when he did not secure a place in grammar school. At his secondary modern he obtained two 'O' levels (English Language and Commercial Studies), and passed four more at a college of further education. It took him three years to obtain 'A' levels in Economics, Accountancy and Business Studies (A, C, and D grades respectively). It was a further surprise (and a demoralising one) when Mr Field received a 2:2 Bachelor's Degree in Economics from his first university. He transferred to a university in Scotland and took 30 months to complete a Masters Degree in Economics. At the time of his arrest he was training as a teacher, and

was only a few weeks short of qualifying. This was not a preferred career choice for Mr Field but he had no other ideas about how to make a living. The nature of the index offence is such that he is unlikely to be approved by educational regulatory authorities to work as a teacher, and he no longer wishes to pursue this option. Instead he has found employment as an assistant editor on a banking website. The task provides him with a challenge, is well suited to his training background in economics, allows him to work semi-independently, is lucrative, and calls on his computer abilities. Mr Field believes that he has found his niche. He sees this post as a long term one.

His medical, psychiatric, forensic, and substance use histories are all unremarkable.

Mr Field has never had any longstanding friendships, something his father has been concerned about. There has been one sexual relationship while at his first university. It lasted several months and according to Mr Field ended by mutual consent. They apparently remained friends, and Mr Field transferred to a Scottish university when she did. He watched and followed her for a year in that setting, but without making her aware of his presence. Gradually he lost interest in this woman.

An obsessive interest in two other young women has been reported. One was a fellow student at the college of further education he attended, and another a commuter on a train he caught regularly. Mr Field found out the latter's place of

work and telephoned her there but she was unable to take his call. Nothing further happened.

Mr Field told me that he was not widely experienced in an interactive sexual sense but he has used a variety of pornography to masturbate three or four times a week. Some of this material depicted sado-masochistic practices where women were tied up and whipped. He enjoyed this as although he knew the participants were acting, the women seemed vulnerable and under the control of men. Sources of pornography were the internet, newsagents and sex shops in Amsterdam when he visited there once. Explicit material was most satisfying for him, and the majority of it in his possession allegedly depicted consensual heterosexual activity between adults. It was all confiscated by the police as part of the investigation. No paedophilic material was reported.

Masturbatory fantasies involving Mr Field's victim were few, romantic, and consensual for the most part. He did think about kidnapping her and raping her, but it was important for him to believe that she would experience this as a kind of sex game played out by two people in love who were trying new experiences together. He liked to imagine himself as part of her life, as a saviour and protector, and in therapy struggled to understand how she could view his letters and gifts as aggressive rather than signs of his love.

7.7 Negotiating a contract and therapeutic aims

In any therapeutic endeavour there is an implicit contract between the therapist and the client, wherein the latter receives approval for achieving a variety of sub goals on the way towards eventual termination of contact (Rimm & Masters, 1974). Deals may be more explicit as in attempts at controlling obesity using contractual procedures where deposited money is returned as each pound of weight is lost (Harris & Bruner, 1971). With both of these approaches there is an understanding that counsellor and client are involved in a mutually co-operative attempt to comprehend the world better, share each other's perspectives, and resolve some agreed upon problem. In Mr Field's case however, this was not so. He saw his behaviour as non-threatening, attempted to impress a woman he believed he was in love with, and that both she and the authorities had reacted out of all proportion to what he had done. Furthermore, he was mildly traumatised at being put in prison, and being an intelligent man did not trust those charged with his care and treatment to respect his privacy. When we first met he was polite but guarded in his responses, and I took some pains to explain the limits of confidentiality we would operate within. In certain respects this differs from how most psychotherapeutic engagements begin, and more closely resembles phase one of a police interview where the purpose of the meeting is stated, participants are introduced to each other, and legal requirements fulfilled (Gudjonsson, 1992). It is important to address these issues at the outset, as one cannot ignore the atmosphere of hostility and surveillance present in any prison setting. To do otherwise would be disrespectful to the client and a dishonest beginning to treatment.

Mr Field had already been sentenced when I first met him, but we both needed to be aware that our exchanges might have an impact on his accommodation in the prison, the work he might be allowed to do, his chances of obtaining early release under the parole system, his living arrangements after discharge, and follow-up supervision by the Probation Service. Such a power imbalance between parties in therapy needs to be acknowledged and not abused. A central part of the counselling relationship is to develop a collaborative alliance (Nelson-Jones, 1995b). This is a standard practice of mine within the first minutes of the initial meeting. I explain who I am, where I am based, the relationship between my service and the prison, who has made the referral, its purpose, what documentation will be involved, my reaction should the client tell me of intent to self-harm or behave violently, and how long we have together today. It is made clear that the client does not have to see me if they do not wish to (even if I have been asked to provide a court, prison, or Probation Service report). This achieves two results. It gives the client some power back in deciding whether to allow a relationship to form between us or not, and it decreases the likelihood of violence to me (and a further blot on their record). I cannot remember the last time somebody terminated the meeting at this point. Usually it leads on to lively discussion of the issues raised.

Mr Field's curiosity was piqued by my introduction, and we were able to agree to meet as regularly as my schedule would allow. The deal was that we would continue as long as we both thought there was something to be gained by it, hoping that control would not be taken away from us by an unexpected transfer to a different prison. The principal focus for our work would be an examination together of some

200 pages of written material Mr Field had sent to the victim between 27/5/98 and 04/06/99 when arrested. The goal was to enhance our understanding of the offending behaviour. We agreed to look for themes, irrational beliefs, faulty reasoning styles, and inadequate evidence Mr Field had used to sustain his inappropriate behaviour. If progress allowed we would consider relapse prevention strategies. The first therapeutic appointment proper was set, and I explained the underlying principles of REBC.

7.8 Key content issues

As mentioned earlier in this report, working with someone like Mr Field in the forensic context differs significantly from the usual counsellor/client relationship. Along with the consideration of Mr Field's needs, thought (and sometimes action) must be directed to issues of how his behaviour fits within the law, and the rights of people living in his community to be protected against antisocial acts. Evidence from the United States suggests that up to 25% of stalking cases culminate in significant violence (Harmon et al, 1995; Meloy and Gothard, 1995).

Thinking principally of Mr Field, the assessment process highlighted a set of irrational beliefs he held about the victim. These were not the result of mental illness and seemed the sort of material which might form the basis for an REBC approach. People usually act on the basis of thoughts and emotions, and their actions influence how they think and feel (Ellis, 1991). Mr Field's settled mental state, academic accomplishments, awareness of his predicament, and vested interest in change to minimise the chances of relapse (and impress custodial staff with his safety) further supported this choice of therapeutic model. It also helped that my theoretical orientation is cognitive-behavioural psychology.

In retrospect this seemed a sound point to start from, particularly as a review of the stalking literature has since stated, "there are no reports of the development of any specific treatment programmes, either for stalkers or for their victims", (Kamphuis & Emmelkamp, 2000, Page 206).

Moving on to public protection issues, it was clear to Mr Field and I from the outset that our work together would need to be viewed as a series of risk assessments, and that I would be attending meetings where his personal growth, change, and potential for future offending would be discussed. They would include staff from the Probation Service, police, prison, public protection unit, and housing organisations. At such times I would have a dual role as therapist giving feed back on Mr Field's progress, and scientist-practitioner attempting to inform and moderate the reactions of other participants. Managing a stalker can be very anxiety provoking, especially for those with no clinical background and who have only read reports in the popular press about what such people are capable of. The most extreme cases get over reported and then any one remotely connected with a stalker worries about what might happen and how that could reflect on their professional competence. Not all stalkers are the same, nor do they pose the same level of risk to others. In their article "Study of Stalkers", Mullen et al (1999) recognise five types: rejected, intimacy seeking, incompetent, resentful, and predatory. From what I knew of Mr Field he most closely fitted the category of intimacy seeking. This typology does not feature highly amongst stalkers who go on to commit contact offences, and indeed the factors which do predict violent offending (i.e. previous convictions, substance abuse, and previous threats) are absent in Mr Field's case (Mullen et al, 1999). Nevertheless, a significant number of forensic stalking cases culminate in serious violence (Harmon et al, 1995; Meloy & Gothard, 1995), so it is reasonable to expect that members of the multidisciplinary team working with such offenders should communicate regularly with each other, debate opinions vigorously, and collect as

much empirical data as possible when attempting to manage the individual and whatever risk they might pose.

7.9 Therapeutic process

I first met Mr Field on 20/01/00 in prison, and our last appointment was on the 23/04/01 in the community. Over that fifteen month period we had sixteen sessions, the first six while Mr Field was in prison, and the last nine as an outpatient living independently in rented accommodation but under supervision by the Probation Service. It is my practice to document clinical work in considerable detail, and there are on file thirty pages of reports and letters which I have dictated about our involvement. All of this has been shared with Mr Field's Probation Officer, and forms the basis for an REBC analysis of our exchanges. This will be done using the pivotal points of Goals, Activating events, Beliefs, Consequences, Disputation, and the development of Effective new philosophies as advocated by Ellis (1991).

These are presented in chronological order within each subsection to allow the reader to see how therapy progressed over time.

7.9.1 Goals of the stalking behaviour

Forming a relationship with the victim.

Showing her how he feels about her (see appendices 1 and 2).

Making her love him.

Having someone important in his life.

Having someone he can devote himself to (obsessively).

Having someone he can share with.

Being in control of the relationship.

7.9.2 Activating events

These are defined in Nelson-Jones (1995a) as things which happen (or are done) and that become linked in a persons mind with how they feel at C (i.e Consequences). In Mr Field's case they include;

Sending the victim letters, postcards, gifts, and sweets.

Watching her on television.

Recording her shows to watch again.

Clandestine visits to the television studio (see appendices 3, 4 and 5).

The victim's non-verbal communication, which Mr Field misread and over-interpreted.

Arrest, trial, and imprisonment.

7.9.3 Beliefs

I can make her love me by persistence.

My overtures will make her curious, interested, and eventually agreeable to a meeting with me.

I am intelligent, desirable, sexually proficient enough to sustain a relationship with a celebrity.

She is not really frightened of me.

I can reassure her of her safety by making explicit references to serious crimes committed by other stalkers (see appendix 4, paragraph 2). This way she will know that I would never behave like that.

I know I am not going to harm her so she should know that too.

I know what she thinks and can speak on her behalf in therapy.

I am the victim here. I got a 30-month prison sentence while she became more famous as a result of my attention.

My offence is not as serious as most other stalkers.

She looks OK when I see her on television giving interviews about this case, so I can't have had a bad effect on her.

People don't need to worry about what I might do.

Any negative effects of my attention to her are far outweighed by my love for her.

My efforts to reassure her of her safety include:

(a) a list of advantages and disadvantages of having a stalker (see appendix 6)

(b) the golden rule of all stalkers (see appendix 7)

I am certain I will never reoffend or pose a serious risk to anyone.

I am incapable of violence towards others, especially women.

My story about the two brothers who suffered from mental illness and who killed a woman in the woods was a random choice as I was running out of things to put in that letter. I don't know why she inferred I was planning to waylay her in the woods and harm her.

The Grim Reaper postcard was chosen at random, and what I wrote on it was never intended to frighten her, no matter what the judge thought (see appendix 8).

7.9.4 Consequences

I feel good knowing we have this special relationship.

My life has more purpose to it.

I'm special, different, because I have a celebrity "girlfriend".

I have something to look forward to – the eventual consummation of our relationship.

I feel OK when I think about her because despite what everybody says she doesn't look upset to me. After all, I don't want to hurt the one I love.

Any more stalking behaviour would be dealt with severely by the court.

7.9.5 Disputation

REBC theory was explained to Mr Field. Direct, confrontational challenging of the evidence he used to support his beliefs and the viability of his goals occurred regularly. This is how most REBC practitioners work (Nelson-Jones,1995a).

Scientific questioning and the search for proof was appropriate in view of Mr Fields academic background. Discussion was often heated, with opinions expressed passionately, but there was no sense of personal danger for me as a therapist. I saw this attribute of Mr Field's to enter into lively debate without crossing the boundary into personal abuse as an encouraging sign of psychological adaptability.

Initially his resistance to rational disputation was strong. There were many "perhaps", "maybe", and "yes, but..." responses. I reflected back to Mr Field that in my understanding these utterances were the equivalent of saying "no", and he agreed.

Mr Field was encouraged to consider a range of possible explanations for phenomena rather than grasping one and distorting the data to fit it.

Insight and the internalising of more rational philosophies was slow to develop. My report on session four notes that Mr Field was rigidly argumentative with no signs that he was modifying his thinking style in any significant or healthy way. The risk he posed of further antisocial conduct was not seen as decreasing.

We began to consider possible costs to stalking victims (e.g. curtailed behaviour, change in self-image, quality of life decrements, curbed spontaneity, etc). Pathe & Mullen found in their 1997 study that on the basis of self-reports, 37% of their sample qualified for a diagnosis of Post Traumatic Stress Disorder. These repercussions were highlighted by using extracts from his letters to the victim. This had an impact on Mr Field. He showed non-verbal signs of embarrassment (e.g. breaking eye contact, grimacing, gritting his teeth) when I read out these passages. Although difficult, I was able to keep him focused on these topics, at which times his arguments gradually dried up. However, he used distracting techniques (i.e. linking the idea under discussion with another) and once safely distanced from the words he

had written his confidence returned, and he gave the impression of being just as convinced as before that his actions had been harmless.

Mr Field felt that he was no risk to anyone. I disputed this (on 3/4/00) by referring to the pursuit of his German ex-girlfriend in Scotland for a year, and the lengthy, ambiguous correspondence he had entered into with the victim. My point was that nobody could be sure that his risk taking behaving might not escalate in future in response to growing awareness of frustration and rejection.

I drew comparisons between Mr Field's circumstances, attitudes, and personal characteristics and those of other people who had demonstrated obsessional love, and formed the conclusion that there was enough overlap to raise concerns amongst those dealing with his case (Kamphuis & Emmelkemp, 2000). In the United States it is estimated that up to one quarter of forensic stalking episodes culminate in significant violence (Harmon et al 1995; Meloy & Gothard, 1995). Mr Field could not tolerate this view point, and maintained for some time that he was incapable of violence towards others, especially women. I was eventually able to convince him that neither he nor anybody else that I have ever met was able to predict the future with 100% accuracy, and he reluctantly conceded that, although in his mind the balance of probabilities favoured that he would not commit a violent act, he could not rule one out altogether.

The nature of the relationship between Mr Field and the victim needed clarification. On Page 126 of the Crown Prosecution Service evidence bundle Mr Field wrote to the victim that he was in a university computer laboratory session when he read an e-mail the woman next to him was about to send. She was writing to a man she loved but who did not feel the same way about her, and was threatening to commit suicide when she was next in his flat. Mr Field wrote to the victim, "I compared the mess her life was in to our relationship. It was very apt. We're really happy together, aren't we?" When we discussed this he changed his mind about his earlier formulation, and agreed with me that the woman at the computer seemed to have many similarities with him in that they were both involved in instances of unrequited love. He seemed a little embarrassed about this realisation, and it was one we came back to several times as therapy progressed.

7.9.6 Effective new philosophies

My behaviour was ambiguous. The victim and the court believed that. I shall strive to avoid ambiguous conduct in the future so the police will leave me alone.

Once bitten, twice shy. I don't want anything to interfere with my freedom in the future. I hate prison.

This is not the start of a criminal career, it is a serious lapse of judgement which I shall avoid in future. My history (family, friends, medical, psychiatric, use of intoxicants) supports the belief that I can avoid further legal trouble.

My behaviour could be seen as "addictive" or "dependent". Abstinence is a good relapse prevention strategy.

The only way to know what people think, feel, and believe is to ask them. I cannot guess with any sense of certainty.

Continuing to watch my victim on television poses a risk of reawakening an unhealthy interest in her. This could have negative consequences for both of us.

I can immerse myself in my career to satisfy my need for obsessional involvement.

My new job has the potential to become a career. That's where I will channel my energies rather than dwelling on stalking.

7.10 Reactions to the therapeutic process

The major difficulty in working with Mr Field was challenging his irrational beliefs and protective rationalisations about the offending behaviour. He had become so obsessively involved in his fantasies about the victim that he had withdrawn from social contact, thus preventing the chance of any rational feed back about what he was doing. It was a novel experience for him to hear and think about my perspective as a therapist. Rapport was not difficult to establish, however, and once his intellectual curiosity was aroused he found the process interesting enough to continue with.

In between sessions I was able to use meetings with my clinical supervisor, line manager, and colleagues to review my formulation of the case. In fact, with Mr Field's permission I used his circumstances as a basis for a case presentation to Belgian nursing students and their tutors who came to spend a day with our service. The overall theme of the speakers was risk assessment and management. Copies of the 23 overheads used were sent to Mr Field, and going through this process was helpful in focusing my thinking about what we were doing together.

It was clear from the outset that mental illness was not an issue with Mr Field despite his sometimes unusual ideas, and protracted involvement with him reinforced this. There was no evidence to suggest a need for psychotropic medication or elaborate consideration of diagnostic categories. With each session that was completed the choice of REBC as a therapeutic modality seemed more sensible, and neither Mr

Field nor I saw any reasons to change it. Eventually, less and less challenge was needed as Mr Field presented fewer irrational beliefs about himself, the victim, relationships and stalking. He continued to hold strong views about the legal system and his place in it, but these were based on objective reality and were not significantly divergent from those of many other prison inmates and parolees.

7.11 The therapeutic ending

While Mr Field was a sentenced prisoner and then a parolee on probation who could be returned to complete his sentence in prison if his supervising officer felt certain conditions were not being met, there were obvious pressures on him to attend sessions with me. He and I discussed these only briefly, as he was more curious about therapy than feeling coerced in to it. We both agreed that the initial resistance to change was more likely a function of his passionately held belief systems rather than anger at being referred to a psychologist. Our 16 sessions over 15 months gave us enough time to feel that we had explored the relevant issues in sufficient depth to be satisfied that nothing major had been overlooked. Our last meeting took place four days before his probation supervision period ended. From then on there was no statutory obligation for Mr Field to continue seeing me, but in any case we both felt that further joint work was unnecessary. I reviewed relapse prevention strategies with him and emphasised the importance of contacting his General Practitioner for a counselling referral if he felt worried about the potential development of unhealthy preoccupations. It was made clear that this could be done without any chance of compromising his freedom. The legally sanctioned punishment period for his offence had expired, and seeing a therapist for related concerns would not automatically result in police or court action.

Mr Field was satisfied that he had sufficient protective measures in place to feel safe about the future. His parents continued to be supportive, the accommodation he lived in was comfortable, his new job kept him stimulated without feeling overwhelmed, for the first time in his life he had a reasonable salary coming in, there

were no urges or cravings to pursue his victim, nobody else had taken her place in his thoughts, there was no abuse of intoxicating substances, he now had the psychological tools to challenge irrational thinking, and his mental health was settled.

With regard to the latter, although there were no signs of diagnosable mental health problems at interview or from the way Mr Field was coping in the community, I felt it prudent to administer the Minnesota Multi-Phasic Personality Inventory (2nd edition) as a symptom scan. It consists of 567 items to which the respondent answers true or false. The pattern of scores produced can give the clinician an indication of diagnostic categories which might be applicable. Mr Field had completed this while on remand as part of his assessment by Dr Frank Vingoe so there was a baseline to compare his responses with. In both cases the validity scales (i.e. L, F, and K, or lie, infrequency, and correction respectively) were well within acceptable limits, so it was reasonable to draw inferences from Mr Field's answers to the clinical scale items. At the initial assessment there were slight elevations on 2 clinical scales (Pt or psychasthenia and Sc or schizophrenia). The conclusion was that these scores suggested the presence of anxiety, agitation, apprehension, and social isolation (Butcher, 1969), hardly surprising when one takes into account that Mr Field was in prison awaiting trial for a serious offence with a potentially lengthy sentence attached to it. At the most recent evaluation none of the clinical scales came even close to exceeding the accepted limits, and the profile was unremarkable. This was in keeping with Mr Field's consistent presentation over the time I had known him.

There was no compelling reason to believe that any practice effects may have contaminated the second lot of results. The two evaluations were more than 15 months apart, the validity scales were within normal limits, there were fewer external

pressures on Mr Field at the second testing to try and present himself in the best possible light, and I thought it safe to conclude that mental illness could be ruled out.

This finding has an impact on the sort of arrangements to be made about follow up and liaison with other professionals. With regard to the former I feel that nothing more profound than advice to consult his General Practitioner about potential worries was necessary. The latter had been an integral part of everything Mr Field and I did together. There were 3 dangerous persons conferences convened about him in 2000 (i.e. 28/6;30/8;1/11). The first two took place while he was in prison, but the last one occurred after his release into the community. Representatives were present from the Police, County Probation Service, Prison Probation Department, Area Forensic Psychiatry Service, the Public Protection Unit, and local housing organisations. At each meeting Mr Field's current status was considered, the risks he posed were reviewed, and action plans formulated. The victim was informed of Mr Field's release date, her employer was contacted about the need for vigilance and caution with regard to the behaviour of fans, accommodation options were gone over, probation license conditions explained (for all concerned), police surveillance deemed unnecessary, communication channels clarified, and outpatient treatment agreed.

As the two people most directly involved with Mr Field the supervising probation officer and I were in frequent contact. I wrote to her with the details of every outpatient appointment I had with him. As his licence came closer to expiration it became clear to me that neither the risk Mr Field posed nor his mental state made it necessary for him to have a network of statutory agencies overseeing his activities.

The time was right for him to end supervision, terminate therapy, and move on with his life.

7.12 Reflections

Working with Mr Field was my first professional exposure to stalking behaviour. It gave me the opportunity to learn more about it from the research literature and I was surprised to discover that forensic psychiatry has given scant attention to this phenomenon (Kamphuis and Emmelkamp, 2000). There is no consensus about the exact description of stalking, although a working operational definition already included in this report can be found in an article by Meloy and Gothard (1995). It is a behaviour which pervades the life of the offender, with 25% of perpetrators pursuing their victims for between two and five years (Tjaden and Thoeness, 1997).

Criminal stalkers merely reflect the tip of the iceberg: only 50% are reported to the police, of which 25% lead to an arrest, and only 12% result in prosecution (Tjaden and Thoeness, 1997).

Despite considerable effort the current body of evidence is insufficient for the accurate predication of violent behaviour (including murder) amongst stalkers (Dietz et al, 1991 a,b; Meloy, 1997; Mullen et al, 1999). However, some stable risk factors have been identified such as a history of (domestic) violence, background of psychiatric problems, antisocial personality disorder, and a criminal record (Kamphuis and Emmelkamp, 2000).

The victims of stalkers may react by becoming more cautious, suspicious, anxious, and aggressive (Hall, 1998), or by making significant changes in their social and professional lives such as adjusting daily routines, getting an unlisted telephone

number, buying home security systems, changing jobs, decreasing productivity, or even moving to a new neighbourhood (Pathé and Mullen, 1997).

Taking all of the above into consideration it is clear that stalking behaviour is not merely harmless and eccentric infatuation, but has the potential for serious consequences for all involved. This is why so much time and energy were invested in the assessment and management of Mr Field's case. Although we could all see there were issues he needed to deal with in order to stay safe, there was no unanimous agreement about how this should be done. The research literature was of little help. Enquiries into the treatment of stalkers are notably absent, and there are no clear guidelines (Kamphuis and Emmelkamp, 2000). Some favoured a package of conditions so water tight that Mr Field could not deviate from socially acceptable behaviour, but this was of course impossible. My preference was for a psychological approach to his difficulties, with a clearly defined set of external constraints arrived at through joint discussion, and that were feasible enough so that we would not be setting him up to fail. Although I knew virtually nothing about stalking when I started with Mr Field, I did have a sound grasp of REBC theory and application. Working from first principles during the assessment phase allowed me to evaluate his needs against my abilities, and see that a therapeutic alliance was possible. Managing the conflict between a traditional therapeutic approach centred on the client and at the same time considering public protection issues was a challenge, but one I believe both Mr Field and I were able to meet. We never fell out, there were no sulky silences, our sessions were full of productive dialogue, and by the time therapy concluded we were both more aware of how his problematic behaviour had come about, and what could be done to minimise the chances of

recidivism. At the end we both felt there were grounds for optimism, and that Mr Field was better equipped to avoid further trouble with the law.

It is perhaps worth commenting on the shift in attitude Mr Field underwent over the course of therapy. At the beginning he was prepared to see me for a variety of reasons. Those we discussed were intellectual curiosity about what counselling might involve, a fear of being seen as unco-operative by prison authorities, and a desire to learn if the court decision about him was based on reality. Perhaps they were wrong to view him as a danger to his victim, and he could prove to me that he was really someone special, someone different from most people convicted of stalking-related crimes? As our sessions together progressed it became more difficult for Mr Field to support his irrational beliefs about the offending behaviour, and he gradually moved away from a utilitarian perspective to a more genuinely therapeutic one. Our work together took place over a long enough time period for him to consolidate his gains as he went, and not to be overwhelmed by the speed of confrontation and change.

I do not view myself as an expert on stalking, but take heart from the proof this case has given me that a scientist-practitioner with a generic training base can apply first principles to new types of referrals and make a contribution to understanding, treating, and managing them.

Section C References

Butcher, J.N., (1969), MMPI Research, Development and Clinical Applications. New York, McGraw-Hill.

Dietz, P.E., Mathews, D.B., Van Duyne, C., et al (1991a), Threatening or otherwise inappropriate letters to Hollywood celebrities, Journal of Forensic Sciences, Vol. 36, pp.185-209.

Dietz, P.E., Martell, D., et al (1991b), Threatening or otherwise inappropriate letters to members of the United States Congress, Journal of Forensic Sciences, Vol. 36, pp.1445-1468.

Ellis, A., (1962), Reason and Emotion in Psychotherapy. Secaucus, New Jersey : Citadel.

Ellis, A., (1991), The revised ABCs of rational-emotive therapy (RET), Journal of Rational-Emotive and Cognitive Behaviour Therapy, Vol. 9, pp.139-172.

Gudjonsson, G.H., (1992), The Psychology of Interrogations, Confessions and Testimony. John Wiley and Sons Ltd, Chichester, England.

Hall, D. M., (1998), The Victims of Stalking. In The Psychology Of Stalking : Clinical And Forensic Perspectives (ed. J.R. Meloy), pp.113-137, San Diego, CA : Academic Press.

Harmon, R., Rosner, R., and Owens, H., (1995), Obsessional harassment and erotomania in a criminal court population, Journal of Forensic Sciences, Vol. 40, pp.188-196.

Harris, M.B., and Bruner, C.G., (1971), A comparison of a self-control and a contract procedure for weight control, Behaviour Research and Therapy, Vol. 9, pp.347-354.

ICD-10 Classification Of Mental And Behavioural Disorders; Diagnostic Criteria For Research, (1993), World Health Organisation, Geneva.

Kamphuis, J.H., and Emmelkamp, P.M.G., (2000), Stalking – a contemporary challenge for forensic and clinical psychiatry, British Journal of Psychiatry, Vol. 176, pp.206-209.

Meloy, J.R., (1997), The clinical risk management of stalking : "someone is watching over me ...", American Journal of Psychotherapy, Vol.51, pp.174-184.

Meloy, J.R., and Gothard, S., (1995), Demographic and clinical comparison of obsessional followers and offenders with mental disorders, American Journal of Psychiatry, Vol. 152, pp.258-263.

Mullen, P.E., Pathe, M., Purcell, R., and Stuart, G.W., (1999), Study of stalkers, American Journal of Psychiatry, Vol. 156(8), Aug. 1999, pp.1244-1249.

Nelson-Jones, R., (1995,a), Rational emotive behavioural counselling, Chapter 12, pp.261-292, The Theory and Practice of Counselling, 2nd Edition, 1995, Cassell, London.

Nelson-Jones, R., (1995,b), Multimodal counselling, Chapter 14, pp.324-348, The Theory And Practice Of Counselling, 2nd Edition, 1995, Cassell : London

Pathe, M., and Mullen, P.E., (1997), The impact of stalkers on their victims, British Journal of Psychiatry, Vol. 170, pp.12-17.

Rimm, D.C., and Masters, J.C., (1974), Behaviour Therapy : Techniques And Empirical Findings, Academic Press Inc., New York.

Rogers, C.R., (1951), Client-Centred Therapy; Its Current Practice. Boston : Houghton.

Scott, M.J., and Dryden, W., (1996), The cognitive-behavioural paradigm, Chapter 7, pp.156-179, The Handbook Of Counselling Psychology (1996), edited by Ray Woolfe and Windy Dryden, Sage Publications, London.

Selye, Hans, (1956), The Stress of Life, New York : McGraw-Hill.

Solomon, P., and Patch, V.D., (1974), Handbook Of Psychiatry (3rd Edition), Lange Medical Publications, Los Altos, California.

Tjaden, P., and Thoenness, N., (1997), Stalking In America : Findings From The National Violence Against Women Survey, Denver, CO : Center for Policy Research.

Truax, C.B., (1963), Effective ingredients in psychotherapy : An approach to unravelling the patient-therapist interaction, Journal Of Counselling Psychology, Vol. 10, pp.256-263.

Truax, C.B., (1966), Therapist empathy, warmth, and genuineness and patient personality change in group psychotherapy : A comparison between interaction unit measures, time sample measures and patient perception measures, Journal of Clinical Psychology, Vol. 22, pp.225-229.

Ullmann, L.P., and Krasner, L., (1969), A Psychological Approach To Abnormal Behaviour, Prentice-Hall, Inc., Englewood Cliffs, New Jersey.

Wilkinson, L., (1998), Liaison service provides safety network, Nursing Times, Vol. 94 (37), page 37.

Section D: Critical Review

Chapter 8

A critical review of the psychological literature on the controversy between clinical and actuarial approaches to decision making

8.1 Abstract

When making predictions about human behaviour there are two quite different approaches which can be taken. Evidence may be weighed and evaluated subjectively to reach a decision (the clinical method), or collected and interpreted in a mathematical way, where the probabilities of particular acts occurring are calculated on the basis of population statistics (the actuarial method). Since the 1950s research has been carried out to compare the two , and repeated analysis suggests that actuarial judgements are more likely to be accurate than ones based on accumulated clinical wisdom alone. The evidence for this conclusion is reviewed, along with reasons why there has been so much resistance to accepting it.

8.2 Introduction

One of the first to acknowledge there was a controversy between clinical and actuarial approaches to decision making was Paul Meehl in his 1954 book "Clinical Versus Statistical Prediction: A Theoretical Analysis and a Review of the Evidence". The research findings he surveyed suggested that numerical criterion variables of psychological interest (eg. faculty ratings of graduate students who had just obtained a PhD) could be predicted from other numerical variables (such as scores on the graduate record examination and grade point averages) with greater accuracy by a proper linear model than by the clinical intuition of people presumably skilled enough to make such judgements. From this simple beginning the concept of using a mathematical approach to organise data and make predictions and decisions has been applied to such diverse areas as diagnosis from MMPI scores, the presence or absence of brain dysfunction, who should receive places on university courses, juror decisions, parole violation, response to electroconvulsive therapy, tribunal judgements, interpretation of pathology results, determining delinquency from EEG records, and violent re-offending.

In the USA during the 1960s dangerousness to others became one of the primary criteria for the involuntary hospitalisation of people with mental disorders (Steadman et al, 2000). In the 1970s legal liability was imposed on psychiatrists and psychologists who negligently failed to predict their patients' violence following the Tarassoff case (Webster et al, 1994). In the 1980s the dangerousness standard expanded to statutes

authorising involuntary outpatient treatment (Appelbaum, 1994). In the 1990s, risk assessments of potential for violence became part of the Americans with Disabilities Act, a set of laws designed to protect the employment rights of people with mental health problems, unless those difficulties result in the person being evaluated as posing a direct threat of violence to fellow employees or to customers (Bonnie and Monahan, 1997).

With so much at stake in terms of the protection of the public and the rights of the individual under consideration, it is important that the people making predictive decisions base them on empirically sound models of human behaviour, integrating known data in a scientifically responsible way, rather than relying on imprecise pathways to judgement. This task, along with the publicly contentious results when decision makers get it wrong, has led to a division between those practitioners who favour reliance on clinical experience to arrive at an answer, and others who believe the only ethical way to do this is by putting data about an individual into a mathematical model. In many decision making situations presented in the research literature, statistical prediction has been found to be more accurate than expert clinical judgement (Gottfredson, 1987), but there are drawbacks to this approach. This review will consider limitations inherent in statistical analysis, the results of comparative studies, the variety of mathematical models available, practitioners' resistance to actuarial models, reasons for pursuing the best judgement strategy, and definitions of the terminology used.

8.3 Definition of terms

There are two ways of combining variables from a given data set about an individual or group in order to make predictions (Grove and Meehl, 1996). One is known as the clinical method, in which the decision-maker processes information in their head in order to arrive at a judgement (Dawes et al, 1989). This involves informal contemplation based around past experience, training, and sometimes consultation with others. It is an opaque procedure (Grove and Meehl, 1996), and does not always yield similar results from a similar database.

The other is known as the actuarial approach, where the human judge is eliminated, and conclusions rest solely on empirically established relations between data and the event of interest (Dawes et al, 1989). Variables are combined in a formal, objective procedure to reach reliable decisions which are easily replicated (Grove and Meehl, 1996).

This process is transparent, all the steps are visible, and the outcome is the same no matter how many times it is repeated (Grove and Meehl, 1996). Virtually any type of data is amenable to actuarial interpretation once variables have been operationally defined (Dawes et al, 1989).

There are no readily used synonyms for the clinical decision-making approach, but terms often substituted for "actuarial" are statistical (Dawes et al, 1989), mathematical (Buchanan, 1999) or mechanical/formal/algorithmic (Grove and Meehl, 1996).

The most frequently used ways of reaching a statistical judgement are 2x2 contingency tables, regression analysis, and classification trees.

8.4 Varieties of actuarial models

When researchers first began using empirically obtained data to reach decisions about future conduct, they typically presented their information in the form of contingency tables. Webster et al (1994) demonstrate this in their work on the prediction of violence, where success and failure are considered by measuring prediction against outcome. There are two ways of being wrong and two ways of being right. A prediction of not dangerous (ND) may be confirmed at follow-up as a true negative (TN). Alternatively, a prediction that an individual will act dangerously (D) may prove correct, yielding a true positive (TP).

TABLE SIX : VIOLENCE PREDICTION CONTINGENCIES

		PREDICTION	
		ND	D
OUTCOME	ND	TN Predicted not dangerous and shown at follow-up to be not dangerous	FP Predicted to be dangerous, but at follow-up shown to be not dangerous
	D	FN Predicted to be not dangerous, but at follow-up found to be dangerous	TP Predicted to be dangerous, and at follow-up found to be dangerous

When an individual is rated as not dangerous, but contrary to expectation commits some kind of violent act, a false negative (FN) decision has been made. The chance of a false negative prediction leading to injury or death tempers the judgement of

clinicians, mental health review tribunals, parole boards, probation officers, and members of the judiciary. They run the risk of not only a troubled conscience, but also professional and perhaps legal sanction.

Researchers became dissatisfied with the apparent simplicity of comparing prediction with outcome in a 2x2 tabular format, and began to use the regression method in an attempt to make more sophisticated decisions that better matched the complexity of the task. With this approach weights (or levels of importance) are assigned to variables known to correlate with the phenomenon under study, and the sum of these weights is then used to estimate the likelihood of some outcome (Buchanan, 1999). This sum is then considered against already calculated cut-off scores to evaluate its importance (Monahan, 1997).

Reliance on the regression method in making predictions has been challenged by the development of a classification tree approach (Steadman et al, 2000). This has been applied specifically to the estimation of propensity for violence. It is an interactive and contingent model that allows many different combinations of risk factors to arrive at a decision. Based on a sequence established by the classification tree, a first question is asked of all persons being assessed. Depending on the response, or the answer found through checking the patient's file, one or another second question is posed, and so on until a classification of high or low risk is reached. Steadman et al (2000) found that for cases falling between these two thresholds, the aggregate degree of risk posed could not be distinguished from the base rate of violence for the sample as a whole. By

focusing actuarial attention on cases at the more extreme ends of the risk continuum rather than across the entire range, gains in predictive accuracy are possible (Menzies, Webster and Sepejak, 1985; McNiel, Sandburg and Binder, 1998).

All of the models outlined have had some part to play in evaluating the relative efficacy of clinical and actuarial methods in comparative studies.

8.5 Comparative studies

8.5.1 Diagnostic decisions

A central task in the helping professions is identifying the nature of the problems patients are struggling with. Experienced clinicians often pride themselves on being able to distinguish which category a constellation of symptoms best fits into. In psychiatry and psychology the distinction between neurosis and psychosis can form the basis of heated debate, and practitioners have created their own decision pathways to arrive at a conclusion. Some use formalised tests whilst others favour a semi-structured interview, but most believe their own particular approach yields more accurate and consistent results than any other. The differential decision is important. A diagnosis of psychosis can lead to essential (but risky) treatment.

Studies were carried out by Goldberg, (1965, 1968) to clarify the process of choosing between the two labels. He used scores from the Minnesota Multiphasic Personality Inventory. This is a 567 item list of statements to which the respondent answers true or false. During the scoring process responses are grouped into categories, and the pattern these form can be used to arrive at a decision about which diagnosis best fits the patient. Through statistical analysis of scores on eleven MMPI scales, and examination of psychiatric patients' discharge diagnoses, Goldberg generated a set of decision rules. The most effective one for distinguishing the two conditions is shown in Figure Four.

Figure 5: THE GOLDBERG RULE

Add scores from three scales (L, Pa, Sc)

Subtract scores from two other scales (Hy, Pt)

Sum less than 45 → the patient is diagnosed neurotic

Sum greater than or equal to 45 → the patient is diagnosed psychotic

Goldberg went to seven different settings and obtained 861 new MMPIs. He compared the accuracy of the rule with that of 29 clinicians who all had access to the same material, and who attempted the same distinction. Correct decisions were made by clinicians 62% of the time, an accuracy level that was exceeded by the decision rule.

To test the idea that additional practice might alter the results, Goldberg gave clinicians MMPI training packages consisting of 300 new profiles with the criterion diagnosis written on the back. This provided immediate feedback on the accuracy of decisions. Even after repeated sessions adding up to 4000 practice judgements, none of the clinicians equalled the Goldberg Rule's 70% accuracy rate. By always following the rule every one of them would have done better.

Goldberg used the same 861 MMPI protocols to construct mathematical (linear) models of each of the 29 clinicians to reproduce their decisions as closely as possible (Goldberg, 1970). He analysed the relationship between the information available to the clinician and that person's decisions. If the practitioner weights variables with perfect

consistency or reliability (ie. the same data always lead to the same decision), then the model will always duplicate that person's judgements. In reality, people are not perfectly reliable, so clinician and model sometimes arrive at different conclusions. In cases of disagreement, Goldberg found the models were more often correct than the very people on whom they were based. The most likely explanation is that models are perfectly reliable. There is only one pathway for data to be processed, and no matter how often this is repeated the result must always be the same. With human judges, however, the decision steps are unclear and not reliably uniform, so the outcome is open to influence by what has gone on in the practitioner's life since the last conclusion was reached.

8.5.2 Meta analyses

Grove and Meehl (1996) completed a meta-analysis of the literature comparing clinical with actuarial prediction. They included any study that had been published in English since the 1920s, concerned the prediction of health-related phenomena or human behaviour, and contained a description of the empirical outcomes of at least one human judgement-based prediction and at least one mechanical prediction. They found 136 such studies, which gave 617 distinct comparisons between the two methods. Prediction criteria covered a wide range including medical and mental health diagnosis, prognosis, treatment recommendations, treatment outcomes, personality description, success in training or employment, adjustment to institutional life (prison or military), parole violation, violence, and bankruptcy of firms. The clinicians included psychologists, psychiatrists, social workers, and members of parole boards and

admission committees. Their level of education ranged from high school graduates to qualified medical sub-specialists. Levels of experience varied from none at all to many years of involvement in professional decision making. Mechanical prediction techniques started at the simplest (ie. cutting scores), to sophisticated methods involving advanced statistical processes. Data on which judgements were based went from crude tallies of life history facts to sophisticated medical tests. In the majority of studies clinicians had all the information that was available to the actuaries, plus some other data, while no study gave the actuary more data than the clinician. All predictive accuracy statistics were ultimately converted to study effect size units, an approach supported by Bonta et al (1998) in their meta-analysis. When applying these criteria Grove and Meehl (1996) found 64 studies favoured the actuary, 64 showed roughly equivalent accuracy, and 8 favoured the clinician. The latter group were not concentrated in any one area of prediction, did not over-represent any particular type of clinician, and did not seem to have any obvious characteristics in common. The most plausible explanation for these deviant studies is that they arose from a combination of random sampling errors (8 out of 136 or 5.9% of the total) and the clinicians' information advantage in being provided with more data than the actuarial formula. The authors concluded that the great preponderance of studies either favoured the actuary outright, or indicated equivalent performance. The few exceptions were scattered, and did not form a pocket of predictive excellence in which clinicians could profitably specialise.

Their findings were strengthened by the criteria they used to exclude studies from their pool of data. Research was left out if inadequate descriptions of methodology were

given, if predictions were made on different sets of individuals (thus avoiding contamination by studies on cases that might have been easier to predict), if predictive information was not either the same (or a subset) of the data available to each method, or when results of forecasts could not be quantified as correlations between predictions and outcomes, hit rates, or some similarly functioning statistic.

8.5.3 Juror decisions

Although testimony by mental health professionals is widely used where risk of violence is an issue, relatively little is known about the impact of such information on juror decision making. Guy and Edens (2003) conducted a study addressing the effects of testimony based on three types of risk assessment method (ie. clinical opinion, actuarial assessment, and ratings of psychopathy). Their intent was to measure the effect of each one on jurors' perceptions of the defendant. In a mock sexually violent predator civil commitment trial, 172 undergraduates were presented a case summary that included prosecution and defence expert testimony related to violence risk based on one of these three methods. The hypothesis that a defendant described as a "high risk psychopath" by the prosecution would be judged more severely than a defendant judged as "high risk" was supported for female jurors. The authors concluded that perceptions of dangerousness may be more a function of diagnostic label than of the examiner's statement about the defendant's likelihood of offending (ie. high risk). Counter to prior studies, little support was found for the hypothesis that clinical opinion would be more influential than actuarially based testimony. Limitations of the study were presenting expert witness testimony as a one-page synopsis of the evaluation

procedures used and conclusions reached, the sample size was small (particularly after divisions by gender), it is questionable that results from a university student cohort will generalise to the population at large, and explanations given to participants about the relevant legal standards to use when reaching a decision were simplified.

The differential impact of clinical opinion over actuarial evidence on jurors' decisions was demonstrated in a study by Krauss and Lee (2003). Participants were 114 university undergraduates. The procedure involved a simulated capital sentencing hearing. After reading case details, jurors were asked to rate their beliefs about the defendant's dangerousness, and their confidence in that belief. Expert testimony (either actuarial or clinical) was then presented via videotape, and participants re-assessed dangerousness (and confidence in that rating). They also evaluated the expert's level of scientific knowledge, confidence, credibility, and the influence the expert's testimony had on their decision. Cross-examination of the expert was presented on videotape, and participants rated the same measures they had previous completed. A period of 15 minutes was allowed for group deliberation, and final ratings of defendant dangerousness (and participants' confidence in that belief) were obtained. Although influenced by both types of testimony, mock jurors were significantly more affected by clinical opinion expert testimony in their confidence ratings of the defendant's dangerousness throughout the proceedings, immediately after expert testimony was presented, after cross-examination, and after group deliberation. The opinion that mock jurors, at least, are unable to appreciate the weaknesses of clinically based estimates of dangerousness even after cross-examination and group deliberation was

supported. The authors conceded that none of the results explained why clinical opinion was favoured, but wondered if jurors might distrust statistical explanations of events because of a belief that it is more common for statistics to deceive than for human experience to lie.

Limitations of the study were noted as a non-representative jury pool, a short deliberation period, and the artificial nature of simulated trials, as participants know their decisions will not lead to someone being executed.

8.5.4 Prediction of violent behaviour

One of the most contentious areas where clinical judgement and statistical methods have clashed is the prediction of violent behaviour (Webster et al, 1994). This is of interest not only to protect the public, but also to ensure that individual freedom is restricted only when necessary. Under most circumstances predictions of violence remain untested. The decisions of Courts, review and parole boards, and prison officials are guided by clinical opinion. Any suggestion of potential for violence is likely to result in lengthy detention in prison or hospital, or close supervision after release. Proper scientific study would require that all patients, both those considered dangerous and those assessed as non-dangerous, be released into the community and followed up to determine actual levels of violent behaviour. This would ensure that every patient was given the chance to be violent without externally imposed restrictions limiting that possibility.

The research literature on violent behaviour and its prediction is dominated by studies involving male participants. This is, in part, due to the fact that more men than women commit crimes serious enough to result in legal sanctions. In Scotland, only 6% of the prison population are women (Scotsman, 9 March 2005), a rise from 3% in 1995 (Women Offenders – A Safer Way, 1998). In North America the figure was around 6% in 1998 (Gowdy et al, 1998). Scientists have found it easier to gain access to appropriately sized sample cohorts in male prisons and hospital wards.

A decision of the United States Supreme Court in 1966 gave the opportunity for a landmark study. Johnny Baxstrom served his sentence in a prison hospital as a mentally ill inmate, and when his release date came he was transferred under the American equivalent of the U.K. Mental Health Act (1983) to a hospital in the community. In Baxstrom versus Herold the Supreme Court ruled this procedure was unconstitutional. From this legal precedent nearly 1000 inmates in hospitals for the criminally insane were transferred to public hospitals when it was judged they should have been there in the first place. Eventually about half of these were released into the community.

The Baxstrom ruling allowed the kind of natural experiment that could never have been carried out by social scientists. Objections from ethics committees would have been insurmountable. It became possible for Steadman and Cocozza (1974) to test the positive predictions of dangerousness that had resulted in the inmates' hospitalisation.

The follow-up period was 54 months, and 98 released patients were considered. Steadman and Cocozza found two variables had some relation to arrests. These were age and previous criminal activity. A higher likelihood of being arrested was associated with youthfulness, and extensive criminal history. Using these variables correctly identified 17 patients as high risk, but mis-identified 19 as false positives. Despite a total of 76 correct predictions (59 true negatives and 17 true positives), the authors concluded : *"any enthusiasm for success is tempered by the knowledge that these measures, at their best ... would still mean that any special programme or preventative detention established would for every 100 patients classified be inaccurate for over half of them",* (page 148). Of the 20 arrests, 13 were for non-violent offences. The authors used this knowledge to adopt a more convincing measure of dangerous outcome, classifying as violent the seven subjects arrested for violence and adding to this seven others who were hospitalised as a result of their violent behaviour. For every correct prediction of violence from age and previous criminal history there were more than two false positive errors. Clinicians and administrators assumed patients to be more dangerous than they actually proved to be.

It took only a few years for the Baxstrom findings to be replicated in Pennsylvania. There was a transfer of 586 male patients from Farview Hospital for the criminally insane to a number of public hospitals, and many were eventually discharged into the community. Thornberry and Jacoby (1979) set out to repeat the Steadman and Cocozza study with some refinements. They identified 414 patients who had a chance to offend during the on-average three year follow-up period. From this cohort, 98 (or

24%) were arrested at least once. Considered against crime statistics available at the time, this group committed more violence than did ordinary offenders, a finding which was little different from the base rate in the Steadman and Cocozza (1974) study (ie 20.4%). Working from this crude rate of 24% recidivism, the implicit assumption made by Farview staff that their patients were (as a group) dominated by likely re-offenders was not confirmed, as 76% of them stayed out of trouble.

At admission, Farview patients tended to be around the same age as inmates in federal prisons and state and county mental hospitals. Average stay in hospital was about 14 years, and during this period the majority were not violent. Over half did not have a single incident noted in their files. Thornberry and Jacoby were unable to find variables, singly or in combination, that yielded high predictive accuracy for violent acts. Age was the most powerful single variable. While those under 50 years old were proportionately more apt to be dangerous than those over that dividing line, 82% of this group did not behave dangerously at all. A single prediction that no member of the cohort would behave violently would have been accurate in 76% of cases, allowing one to question the basis for keeping so many people in hospital for such a long time.

The Baxstrom and Farview studies drew attention to issues which have undermined research findings when trying to anticipate what people might do, namely low base rates for violence, and high false positive error in prediction. Given that these men had been detained in hospital for long periods on the grounds that they were dangerous, the rates with which they re-offended violently after release were unexpectedly low. Although it is

generally conceded that rare events are more difficult to forecast than frequent occurrences (Buchanan, 1999), it still seems that those charged with the care of these patients were making such predictions, committing false positive errors possibly in an effort to avoid making false negative ones, and restricting liberty unnecessarily. The findings of several other studies in the 1970s highlighted the same problems (Webster et al, 1994).

Changes in administrative and legal policy have meant that prisoners and patients are considered for release at periodic intervals, and even those who have committed serious violent acts are inclined to achieve eventual discharge into the community (Quinsey, 1980). Such developments will have the effect of increasing base rates of violence, and may eventually decrease its impact on attempts to predict recidivism.

Monahan (1981) recognised the difficulties of low base rates and high false positive error, but still concluded that guarded predictions are not only responsibly offered but may in some cases be vitally necessary.

Acknowledgement of the false positive issue may have prompted clinicians to look for ways to improve their accuracy rates. It is, after all, something of an insult to experienced practitioners that they cannot reliably choose who to detain and who to release without putting members of the public at unreasonable risk of being harmed. A second generation of research into violence prediction began in the 1980s.

Sepejak et al (1983) asked mental health professionals to assess the dangerousness to others of 598 defendants remanded for one-day assessment at the Metropolitan Toronto Forensic Services. They were mostly men between 16-30 years old. Follow-up investigations were restricted to the province of Ontario, and in the case of hospital data to only five of the twenty psychiatric facilities in that area. A further 7% of the sample was excluded from the analyses because their status at follow-up could not be accurately established. Study results were therefore based on two thirds of the original sample of 598 cases.

Two years later hospital, criminal, and correctional records were reviewed, and the outcome scored by external raters in terms of dangerousness to others. Overall psychiatric prediction accuracy was 56%, with a false positive level of 61%, figures reminiscent of those found in 1970 studies. Mindful of the base rate issue, investigators removed 25% of cases for which they had been unable to locate at least a single violent incident over the two-year follow-up period. For the reduced sample the overall level of accuracy improved to about 60%. When psychiatrists said that an individual was unlikely to pose a danger to others in the future they were largely correct. The same held true when they forecast new violence. However, there was still an appreciable number of false negatives (37% of all negative predictions), and even more false positives (44% of all forecasts that a remandee would behave violently). The background variables which yielded significant effects under chi-square analysis were previous offence pattern (violent or not), and a history of imprisonment. These findings are in accord with those of Monahan (1981), and support the widely held belief amongst

clinicians that the best way to predict future behaviour is to consider what a person has done in the past.

This reliance on static, historical variables by contrast with dynamic, clinical ones as useful predictors of violence was further supported by the findings of Klassen and O'Connor (1988). They selected a group of potentially violent male subjects from inpatient admissions to a community mental health centre, and followed them up for six months after discharge. Aware of the concerns about low base rates confounding study outcomes, the authors excluded persons with little or no propensity for violence. It was felt that confining predictive assessment to those who had engaged in or threatened violence would maximise the base rate and increase accuracy. Men were considered violent if they had been or could have been arrested for a violent incident which led to re-hospitalisation. Of the 239 subjects at risk, 193 were not violent while 46 were. Although the level of positive predictive power (ie the percentage of people predicted to be violent who actually were so at follow up) was relatively low at 59%, there were still some variables found to be useful in forecasting outcome. These included the number of arrests for disturbing the peace, arrests for violent crime in the past year, whether there was an assault as part of the presenting problem, whether there was a suicide attempt in the presenting problem, age (negatively correlated), marriage, injury by a sibling before the age of 15, abstract reasoning (negatively correlated), and dissatisfaction with the extended family.

The authors note that their results may be partly attributable to the study design, which used actuarial measures, short-term treatment, and a short-term follow up period. The predictive models they developed were tailor-made for their particular set of data, and this led to the accuracy of classification being over-estimated. The results were not cross-validated, and limited by the pervasive research problem of low base rates of violence.

The difference between static and dynamic variables in relation to violence prediction has been refined by Mills, Kroner, and Hemmati (2003). Participants were 209 male prison inmates sentenced to two years incarceration or longer. The actuarial measure used was the score on the Level of Service Inventory – Revised (LSI-R). Participant files and police records were checked to see if they had re-offended after release from custody. The authors used the term stable when describing variables that have the potential to change. They contrast this with the term dynamic, which suggests a process of ongoing change. Static variables are those that cannot change over time (eg. arrested before age 16). Results showed that stable variables added to the prediction of violent behaviour (and non-violent offending) when combined with static factors. Limitations noted were that classification of items as static or stable variables was solely the judgement of the authors, and that stable variables were not shown to be changeable within the period of the study. Stable variables were measured only once, thus the static-stable dichotomy is more akin to historical versus current or recent functioning.

Despite the search for variables to predict violence, hit rates remained poor. Lidz, Mulvey and Gardner (1993) used a sample of 357 men and women predicted to be violent during emergency psychiatric assessment, and matched them by age, race, sex and admission status with a sample predicted to be non-violent. At six month follow-up they collected outcome data from people who knew the patients, as well as from the subjects themselves and official records. Clinicians' hit rates were better than chance when predicting violence amongst male patients, but they did not do so well with females. Of the men, 45% were anticipated to be violent, whilst 42% actually were. In contrast, it was predicted that only 22% of the female subjects would be violent, while 49% actually were. In underestimating the base rate of violence for women, clinicians made serious prediction errors.

The accumulation of discouraging results prompted researchers to move away from clinical variables assessed by professional judgement as predictors of violence towards more objective measurement. Harris et al (1991) assessed psychopathy using the 20-item Psychopathy Checklist-Revised (PCL-R). Their study sample consisted of 169 adult male mentally disordered offenders. All of these men had spent at least two years in a therapeutic community. They were selected for their youth, intelligence, verbal skills, and highly serious offences. Over half had been found not guilty by reason of insanity.

Although the PCL-R is designed for prospective use by suitably trained professionals, Harris et al (1991) employed it retrospectively by having experienced coders search

through comprehensive clinical files. There is no reason to suggest this variation invalidates scores on the instrument (Webster et al, 1994). Research assistants collected outcome data from the Criminal Code Review Board, the Royal Canadian Mounted Police, National Parole Service for Canada, Provincial Correctional and Parole Systems, and the Provincial Coroner's files. Subjects were categorised as violent offenders if they incurred a new criminal charge for an offence against a person, or were returned to maximum security hospital for violent behaviour against people which could have resulted in criminal charges. There was a relatively high base rate of violence over the mean follow up period of 10 years (67 out of 166, or 40%). This was surprising, especially as these men had participated in one of the most exacting treatment programmes ever launched for this type of patient. It took on average around four and a half years for men to fail.

The PCL-R score alone performed as well, or better, than the best combination of 16 other variables selected by multiple regression. Harris et al (1991) used a cut off score of 25; those scoring under 25 were considered non-psychopathic and unlikely to be violent in the future, while patients scoring 25 or over were grouped as psychopathic and therefore prone to violence. Of the 166 subjects, 130 (or 78%) were correctly assigned on this basis. The level of false positive error was encouragingly low at 23%. Similar results were obtained when the cut off PCL-R score was raised from 25 to the conventional level of 30.

The patient group was unusual in that many had been found insane at trial, and some met the criteria for a diagnosis of schizophrenia. There was also a high incidence of violent criminal behaviour in patients' histories, and a high base rate of subsequent violent recidivism. Study findings might be further limited by arguing that predictions of violent recidivism were at least partially the result of differential response to treatment. Perhaps living in a therapeutic community reduced the rate of violent recidivism for the non-psychopaths, but did not alter the rate for psychopaths.

Because the Harris et al (1991) study included non-psychopaths as well as psychopaths it was possible to determine whether the prediction of violence based on a checklist for psychopathy was equally applicable to other diagnostic groups. There were 48 subjects in this sample who met the DSM3 criteria for schizophrenia. There was no change in correlation between PCL-R score and violent recidivism when these subjects were removed from the analyses, so the investigators concluded that the instrument, as a predictor, has applicability across different diagnostic conditions. Further support for this view was found in a related study of schizophrenic offenders (Rice and Harris, 1992). The PCL-R was the best among several predictors of violent recidivism. It has been widely used in this field, and often compared with other actuarial instruments, for example the Self-Appraisal Questionnaire (SAQ). A five-year follow up study showed the SAQ to be at least as effective as the PCL-R in predicting violent and general recidivism (Loza and Green, 2003). The PCL-R author, Robert Hare, felt obliged to clarify some misconceptions about its role in risk assessment, pointing out that it measures one of the most explanatory and generalisable risk factors related to violent

behaviour, and should not be seen as competing with purpose-built risk instruments (Hemphill and Hare, 2004). The construct of psychopathy has theoretical and practical implications for the criminal justice system. Its utility and explanatory power extend beyond the assessment of risk, and complement instruments such as the HCR-20, and VRAG, providing information which may help clinicians to understand better the offenders and patients with whom they work.

8.5.5 Conclusion

The comparative studies should be viewed as repeated sampling from a wide range of judgement tasks centred around the identification and prediction of human behaviour. It is not possible to have complete knowledge of the elements that constitute this range (Gigerenzer and Goldstein, 1996), so absolute conclusions cannot be reached, and there are serious limitations to the use of actuarial procedures in making predictions. They are far from infallible, sometimes achieving only modest results. Even a specific procedure proved successful in one setting must be periodically re-evaluated within that context, and should not be applied to new areas without assessing whether or not this is appropriate.

Stephen Hart (2003) raised cautions about the use of actuarial risk assessment instruments in sexually violent predator (SVP) proceedings. Interpreting such findings is difficult, because the extent to which probability estimates derived from group data can be applied to individual cases is not known. The accuracy of statistical profiles generated by actuarial measures is limited by small sample sizes on which they are

based, and the small number of risk factors involved in them. He concludes by stating that probability estimates based on group data may not reflect the unique likelihood of a particular behaviour being committed by any individual in that group, in the same way that a mean test score for a group of people may be a result which no one person actually obtains.

However, with a sample of about 100 studies, and the same outcome in almost every case (Dawes et al, 1989), it is reasonable to accept that the actuarial advantage is general rather than exceptional. The answer to the question, "would an actuarial procedure developed for a particular judgement task equal or exceed the clinical method in accuracy?" would have to be "yes" on the basis of the available research. Paul Meehl (1986, page 373) concludes: *"there is no controversy in social science that shows such a large body of qualitatively diverse studies coming out so uniformly ... as this one"*.

Failure to accept a consistent body of scientific evidence over unvalidated personal observation may be seen as a normal human shortcoming, or in the case of professionals who identify themselves as working from empirical foundations, as plainly irrational (Dawes et al, 1989). It might also be considered poor practice and not in the best interest of clinicians or patients, yet it continues to happen. Why is this so? What are some of the reasons people resist the actuarial approach so strongly?

8.6 Resistance to actuarial models

Despite 66 years of consistent research findings in favour of the actuarial method, most professionals continue to use a subjective, clinical judgement approach when making predictive decisions (Grove and Meehl, 1996). An example can be seen in the work of Hilton and Simmons (2001). They examined clinical judgements and autonomous review tribunal decisions to detain forensic patients in a maximum security hospital. Variables considered included the availability of an actuarial risk report at the time of decision making (ie. VRAG score), patient characteristics and history, and clinical presentation over the previous year. Detained and transferred patients did not differ in their actuarial risk of violent recidivism. The best predictor of the 187 tribunal decisions was the senior clinician's testimony, and there was no significant association between the VRAG risk score and clinicians' opinions. The authors concede that the sample was biased in that it contained only long-stay high risk patients, that this may have increased the chances of conservative decisions, that information about non-medical treatments was often absent from the files, and no attempt was made to interview decision makers or examine transcripts of the tribunal panels to learn how they had assimilated information presented to them so as to arrive at a decision.

Some clinicians believe the methods are complementary, and that they use both in a creative mix. The literature, however, shows that in most predictive contexts these two procedures disagree in a sizeable percentage of cases (Dawes et al, 1989). One is, therefore, usually obliged to choose between them.

Complaint has been made that actuarial models do not exist for particular decisions, or the ones that do fail to relate specifically to a practitioner's work setting. However, it cannot be assumed that if actuarial formulae are successful in several clinical populations, and each of them does as well or better than local clinicians, then the formulae will not work well in one's own clinic. The turnover in practice staff, with the more recently trained exposed to different academic and field placements under supervisors with a variety of theoretical and practical orientations, means the subjective decision-making process in each clinician's head is prone to the same validity generalisation concern, and perhaps even more so, than actuarial methods (Grove and Meehl, 1996).

Objections have been raised that statistical predictions are based on aggregates, while clinicians seek to make forecasts for the individual (Buchanan, 1999). The claim is that actuarial figures are irrelevant when dealing with the unique person. Grove and Meehl (1996) give an example from physical medicine to counter this view. Suppose that a person suffering from an illness consulted their doctor, and was advised to have surgery. The patient would most likely enquire about the illness and its treatment, and ask for a professional evaluation of the risks involved to allow a pragmatic decision about how to proceed after the potential costs and benefits have been considered. The doctor would most likely respond with empirically derived data about condition and treatment. They are unlikely to maintain that such information is irrelevant because their patient is a unique human being, not just some statistic. As Bishop Butler (1736) concluded, probability is the guide of life.

Staying with the argument about a unique individual faced with making one specific judgement, Grove and Meehl (1996) illustrate the importance of using statistical knowledge by means of a Russian roulette example. Suppose two revolvers are put on a table, one with five rounds of ammunition and one empty chamber, while the other has five empty chambers and only one cartridge. Any reasonable person who wanted to stay alive would most likely use what they know about basic probability to choose the pistol with five empty chambers. This would give them survival odds of 5:1 while the chances of death are 5:1 with the other selection. It is unlikely that anybody would think the odds make no difference because a choice will be made only once.

Strong views are common in this controversy. Grubin (1997, page 17) argues that the actuarial method is *"of little use"*, while Douglas, Cox, and Webster (1999) state that *"the actuarial method, while useful, is not a panacea"* (page 155). The latter team comment on objections raised by Hart (1998), specifically the unique nature of individually pertinent variables and the validity of using predictive instruments in different settings. Both these criticisms are challenged on the basis of Grove and Meehl's (1996) article.

It has been argued that the really important clinical data used in making decisions about individuals are qualitative, not numerical, so they cannot be used in a formalised way (Grove and Meehl, 1996). Examples are narrative case studies and nurses' file notes. However, anything that can be recorded in a document can be operationally defined,

and its frequency tallied. Encoding data may result in some loss of the subtle aspects of the subject under study, but this raises the question about precisely how much these allegedly unmeasurable features contribute to successful prediction. Some research surveyed in reviews supports the opinion that subjective elements are not of great importance (Dawes, 1971; Sawyer, 1966). By contrast, it has been reported that clinicians make risk assessment decisions on behavioural (observable) cues rather than those supported by the research literature. Elbogen et al (2002) investigated clinicians' perceptions of factors derived from studies of violence prediction. One hundred and thirty four clinicians from four psychiatric facilities completed a survey in which they rated the relevance of risk factors frequently cited in research studies, as well as extra behavioural variables. Results indicated that while participants thought research risk factors were relevant, they considered behavioural variables not usually collected as study data to be significantly more relevant for predicting violence. History of violence was seen as the most important risk factor in all four psychiatric facilities. Limitations noted were that participants rated risk factors in the abstract rather than with a specific patient in mind, the authors did not specify inpatient or outpatient violence with regard to relevance ratings, and being a self-report study may have had an impact on the reliability of preferences stated.

The objection that statistical prediction treats the patient as an object rather than an individual, that it comes between the clinician and the person they work with, is not difficult to challenge. It is fair to expect that scientist/practitioners honouring the doctrine of evidence-based judgement will use means of prediction shown to result in the best

decisions for the people they serve (Dawes, 1979). There is no good reason for this to be passed on in anything other than a respectful, sensitive fashion.

There is a misperception that the actuarial method is overly reliant on mathematics, a discipline seen as orderly, rigid, and deterministic, something the world and those living in it most definitely are not. When events are not wholly predictable, as are most of the things studied in the life sciences, one branch of mathematics (probability theory) and its practical derivative (statistical analysis) shows us how to make decisions in the face of uncertainty (Grove and Meehl, 1996).

Belief can have an impact on violence prediction and its accuracy. Douglas and Ogloff (2003) investigated the impact of confidence on actuarial and structured professional risk assessments. Community violence was the outcome criterion. Raters completed the HCR-20 violence risk assessment scheme for a sample of 100 forensic psychiatric patients. Results showed that actuarial judgements (HCR-20 total scores) and structured professional opinions (ie. low, moderate, or high risk of violence) were substantially more accurate when raters were more confident about their decisions. The authors wondered if raters might have been more confident when HCR-20 scores were either very high or very low, and less confident about intermediate scores. Indices of variability for the low and high confidence groups were compared. Both for HCR-20 scores and final risk ratings, indices of variability were highly comparable between low and high confidence groups. This concern was ruled out as a possible explanation for the findings. Limitations noted were that the study design was a retrospective follow up

rather than a truly prospective one, and that only two raters were used. These were graduate students without the professional qualifications more closely representing decision makers in actual applied settings.

Some have put forward the idea that only naïve clinicians make judgements that are inferior to those reached through actuarial approaches, and that proper feedback about mistakes would allow them to modify their style and beat the formula. Goldberg (1965) did not find this to be so when he gave judges immediate information about the accuracy of their diagnoses from MMPI profiles. After 861 trials the less accurate clinicians moved closer to the better ones, but none did as well as a four variable equally weighted regression equation. Insight never guarantees that what has been learned will be applied reliably, with optimal weights assigned to appropriate predictors. The statistician, on the other hand, can include new items found to be reliable indicators of the characteristic under examination, and the improved equation will produce identical results no matter how many times it is used.

There is one area in which the clinician and statistician are equally handicapped. Both use probability (inexact estimates) when moving towards a decision (Grove and Meehl, 1996). Random sampling variations due to chance influence weights, proportions, and actuarial table tallies in the same way that they confound the clinician's memory. Neither side has an advantage here.

Some hold the view that because the world is constantly changing any statistical formula will quickly become out of date (Grove and Meehl, 1996). There is a double standard implicit here. It is assumed that the changes taking place will rapidly invalidate the formula, but leave the clinician's accuracy unaffected. This is unlikely when one considers the steps involved in making a clinical decision. The judge relies on an informal computational system of their own, with unspecified weights attached to variables on the basis of past training and experience. These are seldom reviewed and upgraded systematically as new evidence is encountered.

From an historical perspective the superiority of formal procedures seems obvious. The achievements of western science after Gallileo are more likely due to the scientific method of compiling objective data in a transparent, replicable way which allows us to see connections, rather than having better brains than Aristotle or Aquinas (Grove and Meehl, 1996). Consistent methods of observing, sampling, recording and calculating allow us to separate knowledge from speculation. This is what the human brain is good at, while it is demonstrably poor at weighting variables subjectively and drawing inferences from the results. As Dawes et al (1989) state: *"a unique capacity to observe is not the same as a unique capacity to predict on the basis of integration of observations"* (page 1671). Grove and Meehl (1996) illustrate this using an everyday example. At the cashier in a supermarket one does not look at the collection of items chosen and guess their cost. The assistant uses a machine to scan individual bar codes, and the data is added electronically to tell us how much we owe. Even the most

vigorous opponent of the actuarial method is unlikely to argue against such an approach.

There may be personal reasons for opposing the actuarial method (Grove and Meehl, 1996). Some might fear unemployment as technological advances objectify the process of risk assessment. The self image of professionals is linked to the value they and those around them place on their technical ability. This concept is threatened by the idea that hours of interview with a patient yield nothing more useful in terms of risk prediction than does checking the file for a handful of static historical pieces of information. The cognitive dissonance arising from this realisation may lead to discomfort, and as people are by and large hedonistic (seeking pleasure and avoiding pain) it is no surprise that they prefer to continue with what they know and trust, regardless of the research findings.

Those opposed to the statistical method may make what Dawes (1994) calls the "vacuum argument". Hoped for evidence in support of the clinical approach to decision making is believed to exist somewhere, while negative evidence that has actually been collected is ignored. Proponents of this idea believe that if only the "right sort" of study were done, then the actuarial supremacy would be reversed. When there are 136 interpretable studies with only 5% deviant and ranging over a wide diversity of predictands, and the observed facts accord with theoretical expectations, then it is reasonable to conclude that actuarial approaches to decision making have a greater likelihood of being accurate (and therefore useful) than clinical ones (Grove and Meehl,

1996). This view is supported by Douglas, Cox and Webster (1999); *"what seems clear is that actuarial prediction methods, which typically lack the biases, idiosyncrasies and imperfections of human clinical judgement, are superior to clinical prediction methods in terms of accuracy. Nevertheless, many clinicians are reticent to accept this superiority"* (page 155). Continued rejection of the evidence supporting actuarial methods may be seen as resistance to scientific discovery (Barber, 1961).

One impediment to the adoption of actuarial decision strategies by clinicians is that many are based on main effects regression analyses, and are not seen to reflect the contingent nature of the clinical assessment process. The use of linear regression models implies that a single solution fits all cases, but clinicians clearly do not believe this (Gigerenzer, 1996). A way around this obstacle, which might allow clinician and actuary to co-exist, has been proposed by Steadman et al (2000). They advocate the use of a classification tree (CT) to make decisions, and have carried out research on the prediction of violence. They believe this method reflects an interactive and contingent model of violence which allows for many different combinations of factors to classify a person as high or low risk. Based on a sequence established by the CT, a first question is asked of all persons being assessed. Depending on the answer given (or that found in the person's file) one or another second question is posed, and so on until a judgement of high or low risk is reached. This contrasts with the regression approach in which a common set of enquiries is made of everyone, with each answer weighted to produce a score used for classification.

Steadman et al (2000) illustrate their ideas empirically using data from the MacArthur Violence Risk Assessment Study (Steadman et al, 1998). This is the largest database yet assembled to investigate violence risk assessment. A sample of 939 patients recently discharged from acute psychiatric units was chosen and followed up after 20 weeks in the community. This has been found to be the period during which the prevalence of violence is likely to be highest (Steadman et al, 1998). A wide range of risk factors culled from theories of violence and of mental disorder, from research literature, and from the experience of clinicians and researchers was used in assessing patients (Kraemer et al, 1997). In all, 134 risk factors were identified from four conceptual domains (personal, historical, situational, and clinical), and only serious acts of violence were considered. Applying a CT approach to the data allowed the identification of 12 contingent risk factors which permitted patients to be sorted into groups with varying likelihoods of behaving violently. Cut-off scores were chosen with reference to the base rate of violence in the sample. This was 18.7%, as that many patients committed at least one violent act during the first 20 weeks following hospital discharge. Any case the CT assigned a predicted probability of violence that was greater than twice the base prevalence (ie greater than 37%) appeared in the "high risk" category, and in the "low risk" group if the estimated chance of violence was less than half the base rate (ie less than 9%). Using this rationale nearly half the patient group could be classified. The data from those who could not was pooled and re-analysed using Chi-squared Automatic Interaction Detector (CHAID) software. This process was carried out four times until no further groups could be classified as high or low risk. The final Iterative Classification Tree (ICT) model could be used to separate 72.6% of cases

into high or low risk (Monahan et al, 2000). However, the authors note this was a clinical study of violence among people hospitalised for mental disorders, not an epidemiological study of violence among people with mental illness in the general population. The extent to which the accuracy of the ICT might generalise to other clinical settings such as forensic hospitals was unknown.

The classification tree may be just the link needed to bridge the gap between clinical and actuarial decision making. It combines aspects of both approaches, and this familiarity might enable clinicians to feel confident enough to try it. They tend to think about the people they evaluate for violence potential as possessing certain dominant characteristics that, depending on what those features are, lead assessors to explore additional aspects believed to be associated with the increased likelihood of violence (Gigerenzer, Todd, and the ABC Research Group, 1999). The classification tree model in clinical use would consist of a series of questions flowing from one to another (depending on each prior answer), as is the case in many common diagnostic tools and in risk assessment interviews. Some of the factors identified by Steadman et al (2000) are difficult to obtain in most treatment or evaluation settings (Elbogen et al, 1998). For example, the diagnosis of psychopathy using the Hare PCL-SV takes considerable time and expertise to achieve. Monahan et al (2000) have therefore developed a version of the classification tree that incorporates only factors already routinely collected, or easily obtainable during an evaluation. This may provide a mechanism whereby actuarial information formally and transparently obtained can usefully influence clinical decisions.

8.7 Conclusion

Clinicians are not being asked to abandon their training, theoretical models, or professional experience when making decisions. What they are being asked is to realise that the research literature on prediction makes it clear they can seldom do as well as the equation, and in the interests of good practice they should set a high threshold for making judgements which attach more significance to clinical acumen on its own than empirically validated population features (Grove and Meehl, 1996).

An important reason for making decisions is to allocate scarce resources in a cost-effective manner. In the area of violence prediction this concerns judgements about who is released from prison or hospital, when that will happen, and what follow-up arrangements need to be made. Each of these aspects involve staff time, a finite commodity and one for which there is usually intense competition. Staff meetings, ward rounds, and case conferences can become focused on impressionistic considerations, allowing for the "construction" of dangerousness (Pfohl, 1978; Webster, Dickens and Addario, 1985). Anecdote adds to subjective evaluation, and it is possible to inflate levels of risk posed by the patients under discussion. Using, for example, the Risk Assessment Guide from the violence prediction scheme (Webster et al, 1994) allows clinical judgement to be anchored in transparent fact, and lowers the chances of allocating resources to cases where they are not needed.

Awareness of the limitations of using actuarial approaches exclusively in risk assessment has led to a consideration of how statistical findings might link with clinical

opinion to manage uncertainty. Practitioners have also shifted from emphasising violence prediction in a dichotomous way, to an understanding that identifying risk factors allows multidisciplinary clinical teams to prepare treatment plans which optimise the balance between a patient's civil rights, the appropriate use of scarce healthcare resources, and protecting the public (Doyle and Dolan, 2002). The concept of "dangerousness" as a present or absent phenomenon has been replaced by the continuum of risk. Assessment may help not only to identify those individuals likely to engage in future violent behaviour, but also highlight factors which might be addressed to minimise potential risk through treatment interventions or management strategies (Moran et al, 2001). Descriptions of what might reasonably be expected to change the chances of violence present a balanced picture of the patient when compared with the normative sample. This more fairly represents the perceived level of risk. Percentile ranks may not be as informative as stating that the patient is at low, moderate, or high risk for recidivism (Moran et al, 2001). The authors see risk assessment as informing treatment while the patient is in hospital, and providing a basis for integration of service provision as the person moves into less restrictive environments. The opinion that clinical judgement supplemented by empirical understanding of risk factors for violence (and how those factors interact) should be the standard of care to aim for is gaining wider acceptance (Webster et al, 2002)

Actuarial procedures are not perfect. Litwack (2002) found eight studies that purported to compare clinical and actuarial assessments of potential for violence, and concluded the result was a draw. Methods found to be successful in a particular setting should be

regularly audited for accuracy, and not applied to other areas without careful consideration of how appropriate this might be (Grove and Meehl, 1996). Monitoring efficiency of predictions allows practitioners to improve the technique in a way that is responsive to the varying features of the population they are working with. These steps are explicit, allow for informed criticism, and are readily available to other scientists wishing to replicate or extend research. This cannot be said of clinical judgement which rests on implicit, vague processes that are often difficult to specify.

The upper limits in our ability to predict what people will do are thrown into sharp contrast by actuarial methods. Knowing the modest results available from even the best approaches can challenge unrealistic faith in our predictive powers and lead to a sobering humility (Dawes et al, 1989). The genuine scientist-practitioner will want to take this into consideration when making judgements that affect the lives of their patients and those around them.

Section D references

Appelbaum, P.S., (1994), Almost a Revolution: Mental Health Law and the Limits of Change, New York: Oxford University Press.

Barber, B., (1961), Resistance by Scientists to Scientific Discovery, Science, Vol.134, pp.596-602.

Barrett, D., (2005), Only jail females who are danger to public, Scotsman.com. News – Law and Order, Wed 9 Mar 2005

Bonnie, R., and Monahan, J., (1997), (Eds), Mental Disorder, Work Disability and the Law, Chicago: University of Chicago Press.

Bonta, J., Hanson, K., and Law, M., (1998), The Prediction of Criminal and Violent Recidivism Among Mentally Disordered Offenders: A Meta-Analysis, Psychological Bulletin, Vol.123, No.2, pp.123-142.

Buchanan, A., (1999), Risk and Dangerousness, Psychological Medicine, Vol.29, pp.465-473.

Butler, J., (1736) The Analogy of Religion. London : Printed by J Jones for George Ewing UK.

Dawes, R.M., (1971), A Case Study of Graduate Admissions : Application of Three Principles of Human Decision Making, American Psychologist, Vol.26, pp.180-188.

Dawes, R.M., (1979), The Robust Beauty of Improper Linear Models in Decision Making, American Psychologist, Vol.34, pp.571-582.

Dawes, R.M., (1994), House of Cards. New York : Free Press, USA.

Dawes, R.M., Faust, D., and Meehl, P.E., (1989), Clinical Versus Actuarial Judgement, Science, Vol.243, pp.1668-1674.

Douglas, K.S., Cox, D.N., and Webster, C.D., (1999), Violence Risk Assessment : Science and Practice, Legal and Criminological Psychology, Vol.4, pp.149-184.

Douglas, K.S., and Ogloff, R.P., (2003), The Impact of Confidence on the Accuracy of Structured Professional and Actuarial Violence Risk Judgements in a Sample of Forensic Psychiatric Patients, Law and Human Behaviour, Vol.27, No.6, pp.573-587

Doyle, M., and Dolan, M., (2002), Violence risk assessment : combining actuarial and clinical information to structure clinical judgements for the formulation and management of risk, Journal of Psychiatric and Mental health Nursing, Vol.9, pp.649-657

Elbogen, E., Osborne, L., Gordon, K., Tomkins, A., Scalora, M., Penrod, S., Monroe-de Vita, M., and Washington, D., (1998), Clinical Considerations in Violence Risk Assessment. Paper presented at the 1998 meeting of the American Psychology-Law Society.

Elbogen, E.B., Mercado, C.C., Scalora, M.J., and Tomkins, A.J., (2002), Perceived Relevance of Factors for Violence Risk Assessment : A Survey of Clinicians, International Journal of Forensic Mental Health, Vol.1, No.1, pp.37-47

Gigerenzer, G., (1996), The Psychology of Good Judgement : Frequency Formats and Simple Algorithms, Medical Decision Making, Vol.16, No.3, pp.273-280.

Gigerenzer, G., and Goldstein, D.G., (1996), Reasoning The Fast and Frugal Way : Models of Bounded Rationality, Psychological Review, Vol.103, No.4, pp.650-669.

Gigerenzer, G., Todd, P., and the ABC Research Group, (1999), Simple Heuristics that Make Us Smart, New York : Oxford University Press, USA.

Goldberg, L.R., (1965), Diagnosticians versus Diagnostic Signs: The Diagnosis of Psychosis versus Neurosis from the MMPI, Psychological Monographs, Vol.79, No.9.

Goldberg, L.R., (1968), Simple Models or Simple Processes? Some Research on Clinical Judgements. American Psychologist, Vol.23, pp.483-496.

Goldberg, L.R., (1970), Man Versus Model of Man: A Rationale Plus Some Evidence for a Method of Improving on Clinical Inferences, Psychological Bulletin, Vol.73, pp.422-432.

Gottfredson, S.D., (1987). Statistical and Actuarial Considerations. In F. Dutile and C. Foust (Eds.), The Prediction of Violence, (pp.71-81), Springfield, Illinois: Charles C. Thomas.

Gowdy, V.B., (1998), Women in Criminal Justice : Executive Summary, U.S. Department of Justice, Office of Justice Programs, www.ojp.usdoj.gov/reports/98Guides/wcjs98/execsumm.htm

Grove, W.M., and Meehl, P.E., (1996), Comparative Efficacy of Informal (Subjective, Impressionistic) and Formal (Mechanical, Algorithmic) Prediction Procedures: The Clinical-Statistical Controversy, Psychology, Public Policy, and Law, Vol.2, No.2, pp.293-323.

Grubin, D., (1997), Predictors of Risk in Serious Sex Offenders, British Journal of Psychiatry, Vol. 170, pp.17-21.

Guy, L.S.., and Edens, J.F., (2003), Juror Decision-Making in a Mock Sexually Violent Predator Trial : Gender Differences in the Impact of Divergent Types of Expert Testimony, Behavioural Sciences and the Law, Vol.21, pp.215-237

Harris, G.T., Rice, M.E., and Cormier, C.A., (1991), Psychopathy and Violent Recidivism, Law and Human Behaviour, Vol.15, pp.625-637.

Hart, S.D., (1998), The Role of Psychopathy in Assessing Risk for Violence : Conceptual and Methodological Issues, Legal and Criminological Psychology, Vol.3, pp.121-137.

Hart, S.D., (2003), Actuarial Risk Assessment : Commentary on Berlin et al, Sexual Abuse : A Journal of Research and Treatment, Vol.15, No.4, pp.383-388

Hemphill, J.F., and Hare, R.D., (2004), Some Misconceptions About The Hare PCL-R And Risk Assessment : A Reply to Gendreau, Goggin, and Smith, Criminal Justice and Behaviour, Vol.31, No.2., pp.203-243

Hilton, N.Z., and Simmons, J.L., (2001), The Influence of Actuarial Risk Assessment in Clinical Judgements and Tribunal Decisions about Mentally Disordered Offenders in Maximum Security, Law and Human Behaviour, Vol.25., No.4., pp.393-408

Klassen, D., and O'Connor, W.A., (1988), A Prospective Study of Predictors of Violence in Adult Male Mental Health Admissions, Law and Human Behaviour, Vol.12, pp.143-158.

Kraemer, H., Kazdin, A., Offord, D., Kessler, R., Jensen, P., and Kupfer, D., (1997), Coming to Terms with the Terms of Risk, Archives of General Psychiatry, Vol.54, pp.337-343.

Krauss, D.A., and Lee, D.H., (2003), Deliberating on dangerousness and death : jurors' ability to differentiate between expert actuarial and clinical predictions of dangerousness, International Journal of Law and Psychiatry, Vol.26, pp.113-137

Lidz, C.W., Mulvey, E.P., and Gardner, W., (1993), The Accuracy of Predictions of Violence to Others, Journal of the American Medical Association, Vol.269, pp.1007-1111.

Litwack, T.R., (2002), Some Questions for the Field of Violence Risk Assessment and Forensic Mental Health : Or, "Back to Basics" Revisted, International Journal of Forensic Mental Health, Vol.1, No.2., pp.171-178

Loza, W., and Green, K., (2003), The Self-Appraisal Questionnaire. A Self-Report Measure for Predicting Recidivism Versus Clinician-Administered Measures : A 5 Year Follow-Up Study, Journal of Interpersonal Violence, Vol.18, No.7, pp.781-797

McLeish, H., Fairweather, C., and Skinner, A, (1998), Women Offenders – A Safer Way; A Review of Community Disposals and The Use of Custody for Women Offenders in Scotland, Social Work Services and Prisons Inspectorates for Scotland, Published by the Stationery Office as ISBN 011 495 8815

McNiel, D., Sandberg, D., and Binder, R., (1998), The Relationship Betweeen Confidence and Accuracy in Clinical Assessment of Psychiatric Patients' Potential for Violence, Law and Human Behaviour, Vol.22, pp.655-669.

Meehl, P.E., (1954), Clinical Versus Statistical Prediction, A Theoretical Analysis and a Review of the Evidence, University of Minnesota Press, Minneapolois, MN.

Meehl, P.E., (1986), Causes and Effects of my Disturbing Little Book, Journal of Personality Assessment, Vol.50, pp.370-375.

Menzies, R., Webster, C., and Sepejak, D., (1985), The Dimensions of Dangerousness: Evaluating the Accuracy of Psychometric Predictions of Violence Among Forensic Patients, Law and Human Behaviour, Vol.9, pp.49-70.

Mills, J.F., Kroner, D.G., and Hemmati, T., (2003), Predicting Violent Behaviour Through A Static-Stable Variable Lens, Journal of Interpersonal Violence, Vol.18, No.8, pp.891-904

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

Monahan, J., (1997), Actuarial Support for the Clinical Assessment of Violence Risk, International Review of Psychiatry, Vol.9, pp.167-169.

Monahan, J., Steadman, H.J., Appelbaum, P.S., Robbins, P.C., Mulvey, E.P., Silver, E., Roth, L. H., and Grisso, T., (2000), Developing a Clinically Useful Actuarial Tool for Assessing Violence Risk, British Journal of Psychiatry, Vol.176, pp.312-319.

Moran, M.J., Sweda, M.G., Fragala, M.R., and Sasscer-Burgos, J., (2001), The Clinical Application of Risk Assessment in The Treatment –Planning Process, International Journal of Offender Therapy and Comparative Criminology, Vol.45, No.4, pp.421-435

Pfohl, S.J., (1978), Predicting Dangerousness : The Social Construction of Psychiatric Reality, Lexington, M.A : Lexington Books.

Quinsey, V, L., (1980), The Base Rate Problem and the Prediction of Dangerousness : A Reappraisal, Journal of Psychiatry and Law, Vol.8, pp.329-340.

Rice, M.E., and Harris, G.T., (1992), A Comparison of Criminal Recidivism Among Schizophrenic and Non-Schizophrenic Offenders, International Journal of Law and Psychiatry, Vol.15, pp.397-408.

Sawyer, J., (1966), Measurement and Prediction: Clinical and Statistical, Psychological Bulletin, Vol.66, pp.178-200.

Sepejak, D.S., Menzies, R.J., Webster, C.D., and Jensen, F.A.S., (1983), Clinical Predictions of Dangerousness : Two Year Follow-up of 408 Pretrial Forensic Cases, Bulletin of the American Academy of Psychiatry and the Law, Vol.11, pp.171-181.

Steadman, H., Mulvey, E., Monahan, J., Robbins, P., Appelbaum, P., Grisso, T., Ruth, L., and Silver, E., (1998), Violence by People Discharged from Acute Psychiatric Inpatient Facilities and by Others in the Same Neighbourhoods, Archives of General Psychiatry, Vol.55, pp.393-401.

Steadman, H.J., and Cocozza, J.J., (1974), Careers of the Criminally Insane: Excessive Social Control of Deviance, Lexington, MA : Lexington Books.

Steadman, H.J., Silver, E., Monahan, J., Appelbaum, P.S., Robbins, P.C., Mulvey, E.P., Grisso, T., Roth, L.H., and Banks, S., (2000), A Classification Tree Approach to the Development of Actuarial Violence Risk Assessment Tools, Law and Human Behaviour, Vol.24, No.1, pp.83-100.

Thornberry, T.P., and Jacoby, J.E., (1979)., The Criminally Insane : A Community Follow-up of Mentally Ill Offenders, Chicago, Illinois : University of Chicago Press.

Webster, C.D., Dickens, B.M., and Addario, S.M., (1985), Constructing Dangerousness: Scientific, Legal and Policy Implications, Toronto, Ontario : University of Toronto Centre of Criminology.

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

Webster, C.D., Hucker, S., and Bloom, H., (2002), Transcending The Actuarial Versus Clinical Polemic In Assessing Risk for Violence, Criminal Justice and Behaviour, Vol.29, No.5, pp.659-665