Violence risk assessment: Science and practice

Kevin S. Douglas* and **David N. Cox**

Simon Fraser University, Canada

Christopher D. Webster

Simon Fraser University, University of Toronto (Psychiatry and Criminology) and Earlscourt Child and Family Centre, Toronto, Canada

Within the conceptual framework of the scientist–practitioner model, this paper describes how the science of risk assessment has progressed to a point where it can be of definite assistance to clinicians. Similarly, clinicians can be of marked help to researchers as they play a role in concept definition and as they design and carry out studies to determine the accuracy of their predictions. Risk assessments ought to be carried out in accord with state-of-the-discipline knowledge. Until fairly recently, it may have been argued that the state of knowledge did not provide any sort of reliable or trustworthy direction on violence risk assessment. This position seems no longer tenable. Topic areas discussed include violence and risk assessment generally, the actuarial–clinical prediction debate, the validity of violence risk assessments, predictors of violence, violence risk assessment schemes (empirically validated structured clinical decision-making), communication of risk assessment findings, and implications for training.

In recent years, there has been an encouraging number of well conceived and executed research studies showing that, contrary to earlier opinion, there is every justification for trying to improve the accuracy of predictions of violence and the quality of assessments of violence risk (for a review, see Douglas & Webster, 1999*a*). The authors wish to describe a *process* through which violence risk assessments can be carried out. This process, involving the use of violence risk assessment schemes, calls for collaboration among researchers and clinicians and can, it is to be hoped, benefit some patients, convicted persons, and even young children inclined toward early aggression and violence.

The authors believe that the gap between scientists and practitioners is wide, and is narrowing only slowly, if at all. Routine practice is little influenced by strong scientific findings, and frequently research has minimal links to the actualities of everyday clinical life. The authors' position, however, is that each sphere must

*Requests for reprints should be addressed to Kevin S. Douglas, Simon Fraser University, 8888 University Drive, Burnaby, British Columbia, V5A 1S6, Canada (e-mail: kdouglas@arts.sfu.ca).

inform the other in order to attain a level of sophistication in both the science and practice of violence risk assessment that is commensurate with its importance.

This paper is based on the scientist–practitioner model of training in clinical psychology. In such a model, clinical work is guided as much as possible by research findings. Violence risk assessment entails making clinical decisions that, more often than not, can have important consequences. The well-known dilemma is that persons who would not be violent if they were released may end up being detained, and violent individuals, who really ought to be confined to an institution, are released. Both types of errors are serious, and it is highly desirable to reduce them. The authors' position is that the goal of minimizing errors can be forwarded by adopting violence risk assessment strategies that are based in research— 'empirically validated' violence risk assessments. This paper first lays a conceptual foundation based on the scientist–practitioner model of clinical psychology, and then discusses how it may apply to the contemporary science and practice of violence risk assessment.

Science and practice

The 1949 Boulder Conference established a model for training clinical psychologists which has endured since that time. Simply put, clinical psychologists would be trained both as researchers and as practitioners. It has been described as a 'grand aspiration that rarely has been achieved in individual psychologists' (Stricker & Trierweiler, 1995, p. 995). Goldfried & Wolfe (1996) state that since this conference 'the field has struggled to foster the synergy between therapy, research and practice' (p. 104). They go on to suggest that the differences between practice and research need to be examined very closely, as it is a dysfunctional relationship in which both sides view the other with disdain. Kanfer (1990) acknowledges that some individuals have demonstrated how effective practice is based on a knowledge of science; however, his concern is that there is only minimal support for those who are systematically involved in building the relationship between the two. He states: 'A failure to resolve the "scientist–practitioner crisis" not only would lessen the potential impact of advances in psychological science on practice but would also endanger the credibility of a psychological profession by dismantling its most substantial foundation' (p. 269). Nezu (1996) concludes, like Kanfer, that the debate which exists between practitioners and researchers in psychology could have major negative consequences if it continues, as ultimately both sides will lose. He comments that 'despite the definition of psychology as being the *science* of human and animal behavior, those individuals engaged in the *attainment* of psychological knowledge continue to be at odds with those engaged in the *application* of such knowledge' (p. 161; emphasis in original).

While practitioners criticize researchers for being out of touch with the realities of clinical work, researchers suggest that clinicians who are out of touch with the research literature cannot provide effective treatment. What is of greatest concern is the apparent lack of any willingness to reconcile differences. Hayes (1996), while acknowledging that psychology is still a very young field, describes the 'virtual

disconnection' that still exists between science and practice. Beutler, Williams, Wakefield & Entwistle (1995) believe that it is the body of knowledge behind these two positions which is the basis for conflict. Researchers feel that empirical findings are a necessity for correct clinical practice, and clinicians argue that researchers are insensitive to the practicalities of their needs and fail to demonstrate relevance in the context of clinical concerns. They point out that the scientist–practitioner duality pervades disciplines where the scientific method is the primary source of knowledge. Following a national survey of psychologists, they concluded that clinicians use research findings in their work more than academics give them credit for, while academics are less likely to acknowledge the usefulness of clinical work. They take the further position that 'it should be remembered that scientists are more dependent on the practitioner community for the survival of their findings than practitioners are on scientists' (p. 991).

An extreme position is that taken by Fensterheim & Raw (1996), who believe that integration of research and practice is not possible. They feel that it is a model which does not work, as these are two distinct fields. Their plea is for deintegration, as they feel that integration impedes progress. They argue that the two areas should interact on the basis of mutual respect while remaining independent. Somewhat in contrast to this position is Meehl's (1954) concept of the 'honest clinician', an individual who constantly questions success and seeks to know through maintaining a research-scholarly attitude, if it is possible to do better. As Stricker (1992) points out, the honest clinician derives certainty from empirical evidence and recognizes the responsibility to maintain a scholarly attitude that is intended to provide clients with the best treatment possible on the basis of current knowledge. This position is perhaps most strongly stated by Singer (1990), who believes that a central ethic of clinical practice is that it is based on a current knowledge of the research literature.

This concern is crucial to the present discussion of risk assessment as it raises the issue of ethical practice. In summarizing Singer's concerns, Stricker (1992) states: 'Thus, the position that research drives practice is endorsed, and it is also elevated to a principle of ethics. Not only are practitioners who are not informed by research offering an inferior brand of treatment, not only are they less than honest clinicians, but they are unethical in their practice' (p. 546). Stricker pursues his concerns in a convincing manner while attempting to determine the ethical issues involved. He concludes initially that it is not unethical to practice in the absence of specific evidence, as such certainty is not possible within the limits of the existing body of knowledge. However, he states that 'although it may not be unethical to practice in the absence of knowledge it is unethical to practice in the face of knowledge. We all must labor with the absence of affirmative data, but there is no excuse for ignoring contradictory data. An insistence on relying on overlearned, favored, but invalid approaches is not justifiable' (p. 546). He cautions against rigid adherence to a particular approach, while recognizing that no clinician, no matter how well experienced or educated, can be competent in all approaches. Out of this recognition comes the ethical obligation to refer clients to other practitioners who possess the necessary competence to deal with a specific problem. This concern over the ethical implications of ignoring the scientific method is perhaps stated

most strongly by Meehl (1997) who, in condemning a reliance on clinical experience rather than empirical findings, concludes that: 'As for our many prescientific practices, I believe that if we do not take strong steps to clean up our act, some smart lawyers and sophisticated judges will either discipline or discredit us' (p. 98).

It is in this context that the authors' concern about the current 'state-of-the-discipline' knowledge with respect to violence risk assessment becomes relevant. Borum (1996), in his article on how risk assessment in clinical practice might be improved, cites the statement made by Webster, Douglas, Eaves & Hart (1997*b*) that 'the greatest challenge in what remains of the 1990s is to integrate the almost separate worlds of research on the prediction of violence and the clinical practice of assessment. At present the two domains scarcely intersect' (p. 1). It is a balance in the scientist–practitioner model that is sought, and yet, as the preceding paragraphs have suggested, it is not clear how this balance is to be attained. Within the area of risk assessment the authors have commented that it is this balance that is in urgent need of being achieved (Webster & Cox, 1997). The indications are, as discussed in this article, that the science of risk assessment has progressed to the point where it is important, and perhaps ethically necessary, that this information be made available to, and be used by, clinicians in assisting them make what has been described as one of the most complex predictions in behavioural science and law (Grisso & Appelbaum, 1992).

A worthwhile by-product of this process is that an example may be provided of how a truly workable scientist–practitioner model can be achieved. Clearly it is possible to focus on the failings of the model. This may give rise to a sense of the differences between research and practice as being insurmountable. Yet the present authors' position inclines toward optimism. Research and practice should benefit one another. The synergy described earlier is a real possibility. As Stricker (1992) has pointed out, there are different levels of research use in which a practitioner can engage. The first occurs simply through an awareness of research. The second involves the active consideration of the implications that research has for practice. The third possibility, which is identified as the highest order of use, involves the active integration of research and practice. Recently Sobell (1996), in a discussion of strategies for bridging the gap between science and practice, suggests that real integration has yet to occur. She describes this as an elusive alliance that requires more than mere recognition as an issue. The rest of this article demonstrates the workability of combining the goals of research and practice in the context of risk assessment. The authors have chosen to call this 'empirically validated violence risk assessment'.

Empirically validated violence risk assessment

Definition of violence

Given that violence is at the core of risk assessment, it is perhaps surprising that its definition has received little attention in risk assessment research compared to other facets of this research. The present authors adopt a definition that can be found in recent literature on violence risk assessment (i.e. Boer, Hart, Kropp &

Webster, 1998; Hart, 1998; Lyon, Hart & Webster, in press; Webster *et al.* 1997*b*). Webster *et al.* (1997*b*) define violence as 'actual, attempted, or threatened harm to a person or persons' (p. 24). Included under this statement are acts that can reasonably be considered to harm another person. Threats of harm ought to be non-ambiguous to be counted as violent. Boer *et al.* (1998) define sexual violence as 'actual, attempted, or threatened sexual contact with a person who is nonconsenting or unable to give consent' (p. 9).

Lyon *et al.* (in press) point out that definitions that focus on consequences of actions may be too narrow, and exclude such patently violent acts as firing a gun into a group of people but hitting no one. However, these autors also point out that definitions that include non-physical acts such as threatening may be considered too broad. In the large-scale MacArthur risk assessment project in the United States, researchers created categories of violence depending on seriousness (Steadman *et al.*, 1998). For instance, Steadman *et al.* (1998) define violence as 'battery that resulted in physical injury, sexual assaults, assaultive acts that involved the use of a weapon, or threats made with a weapon in hand' (p. 395). They then define 'other aggressive acts' as 'battery that did not result in physical injury' (p. 395). Other researchers have made a distinction between physical and non-physical violence (Douglas, Ogloff, Nicholls & Grant, in press; McNiel & Binder, 1989, 1994*a*,*b*).

The present authors think that there is merit to the fairly flexible definition of violence used by Boer *et al.* (1998) or Webster *et al.* (1997*b*). It is capable of being divided along relevant dimensions such as seriousness (i.e. Steadman *et al.*, 1998), physical/non-physical (i.e. Douglas *et al.*, in press; McNiel & Binder, 1989, 1994*a*,*b*), sexual/non-sexual (i.e. Boer *et al.*, 1998), or instrumentality (Cornell, Warren, Hawk, Stafford, Oram & Pine, 1996). This approach allows researchers to capture the broad spectrum of violent acts as well as to pare down conceptualizations into meaningful categories, which can themselves be subjected to study. This approach also facilitates the identification of predictors that may be most important for different types or categories of violence.

The authors are mindful of the fact that there are many other important considerations regarding violence, such as its measurement (official records vs. self-report vs. observation; Convit, Isay, Gadioma & Volavka, 1988; Douglas & Hart, 1999; Mulvey, Shaw & Lidz, 1994; Steadman *et al.*, 1998); its setting (institutional vs. community; Douglas & Hart, 1999; McNiel & Binder, 1989); its targets (family vs. strangers vs. professionals; Estroff & Zimmer, 1994; Mulvey & Lidz, 1995); its density (number of acts in a specified time period); and its breadth (number of different types of violent acts). These sorts of concerns can and should be integrated into research projects and into clinicians' considerations, and can be fitted within the flexible definition outlined above.

Definition of violence risk assessment

Hart (1998) has provided a useful definition of violence risk assessment as '*the process of evaluating individuals to (1) characterize the likelihood they will commit acts of violence and (2) develop interventions to manage or reduce that likelihood*' (p. 122; emphasis in

original). There are several important parts to this definition. First, the fact that risk assessment is described as a process denotes a continuing assessment procedure. Reed (1997) makes this point clearly when, in writing of the English Care Programme Approach (CPA; Department of Health, 1990), he states that 'risk assessment is not a one-off event but part of an ongoing process of assessment, review and re-assessment' (p. 5). Secondly, risk assessment requires a characterization of the likelihood of future violence, or a prediction of violence. Thirdly, the process of risk assessment invites consideration of what can or cannot be done to avert further similar violence in the future. This leads to the concept of violence risk management, which is vitally important, though perhaps has not been recognized as such until fairly recently (Heilbrun, 1997; Heilbrun, Philipson, O'Neill, Paninopolous, Strohman & Bowman, 1998; Monahan & Steadman, 1994*b*; Reed, 1997; Snowden, 1997; Webster *et al.*, 1997*b*).

This description of violence risk assessment implies that the task is much broader than the mere prediction of violence. Prediction surely is required, but so too is action that stems from the prediction. Prediction requires that the most accurate estimate of future violence probability is forwarded. Management requires that steps are taken to provide treatment, programming, release plans and other mitigating strategies to reduce the likelihood of violence. As Hart (1998) describes, a prediction of 'high' likelihood of violence must be coupled with efforts to disprove the prediction. That is, to suggest or provide interventions or management strategies that will reduce or prevent violence risk is a necessary component of the risk assessment process. Similarly, Reed (1997) discusses the ethical obligation or responsibility of clinicians to act in their professional capacity to reduce and effectively manage violence risk. Just as risk assessment is conceived of as a process rather than a static event, so too should risk management be considered such a process. Snowden (1997) defines risk management as 'the process of systematically focusing on methods of reducing both the severity and frequency of recognized adverse clinical risks for each individual patient' (p. 33). Several commentators have written about the usefulness of having *a priori* risk management strategies that flow from risk assessments (Boer *et al.*, 1998; Heilbrun *et al.*, 1998; Monahan & Steadman, 1996; Webster *et al.*, 1997*b*).

Clinical vs. actuarial prediction

One of the persisting controversies in the risk assessment field is the clinical vs. actuarial prediction debate. This debate requires some explanation and also some comment regarding how it may fit into the risk assessment approach described in this article (or, conversely, how the risk assessment approach may fit within the clinical–actuarial debate). The clinical–actuarial distinction and debate has close parallels to the larger framework of this article, the scientist–practitioner model. Traditionally, actuarial prediction may be considered to be aligned with the scientist side of the debate, and clinical prediction with the practitioner side.

Meehl (1954) is often credited with bringing the statistical–clinical prediction debate or controversy to light in the field of psychology. Generally, statistical (or

actuarial, or mechanical) prediction 'involves a formal, algorithmic, objective procedure (e.g. equation) to reach the decision' (Grove & Meehl, 1996, p. 293). Clinical prediction 'relies on an informal, "in the head," impressionistic, subjective conclusion, reached (somehow) by a human clinical judge' (Grove & Meehl, 1996, p. 294). It is important to emphasize that clinical and actuarial prediction refers to the *process* of offering predictions, and not to the types of variables used. Clinical variables, as defined herein, refer to dynamic factors that can change. They could as easily be subjected to clinical prediction methods as they could to actuarial ones. Similarly, it actually makes little sense to speak of 'actuarial variables', as the term 'actuarial' refers to process, not content. Often when variables are referred to as actuarial, what is meant is that they are static or historical. However, static variables, again, may be subjected to either clinical or actuarial prediction methods.

As Hart (1998) points out, perhaps the most important advantage of actuarial prediction is that it improves the validity (accuracy) and consistency of predictions. Grove & Meehl (1996) report the results of a meta-analysis of studies that compared clinical vs. actuarial prediction. Of 136 such studies, 8 resulted in greater predictive accuracy for the clinical method, 64 showed more accurate prediction for the actuarial method, and 64 showed no difference.

What seems clear is that actuarial prediction methods, which typically lack the biases, idiosyncrasies and imperfections of human clinical judgment, are superior to clinical prediction methods in terms of accuracy. Nevertheless, many clinicians are reticent to accept this superiority. In fact, some make the argument that the actuarial method is 'of little use' (Grubin, 1997, p. 17). It is argued here, conversely, that the actuarial method, while useful, is not a panacea, and that it has limitations in the risk assessment field. The function of actuarial prediction methods is simply that—prediction. Risk assessment as conceived here is broader than prediction. Prediction is a necessary first step. Yet risk management and prevention are equally necessary steps. Once a person is defined as high risk, it is in everyone's best interest to suggest means by which such risk can be attenuated.

Hart (1998) notes several other potential problems. One of these is that actuarial methods tend to involve a small number of variables that are assumed to have equally useful predictive value across different groups of people in different settings. This point is also made by Grubin (1997), who claims that actuarial methods may miss the individual and rare variables that may be of most interest to the case at hand. For instance, if an actuarial risk assessment tool places a person at low risk, but that person has just issued a serious threat to kill some person, should that person in fact not be considered high risk, despite the actuarial forecast?

Grove & Meehl (1996) anticipate this (and other) criticisms of actuarial prediction, and in fact refer to this particular objection as the 'broken leg case' (p. 307).[1] Grubin (1997) argues that one cannot realistically apply actuarial predictive models, derived on aggregate data, to individual cases. Grove & Meehl

[1]The authors give the example of knowing, through an actuarial prediction, that the probability of a certain professor going to a movie on Friday night is .84. However, the professor has a broken leg, making the trip to a movie very unlikely, and broken limbs were not taken into account in the actuarial prediction model. What should the decision be?

(1996) put forth that, though such cases give rise to an important dilemma for the actuary, such cases are also relatively infrequent. In the hypothetical example, the broken leg is a clear objective fact with equally clear implications. As such, countervailing the actuarial equation may be warranted. The danger of ascribing 'broken leg' status to large numbers of predictions is the application of non-optimal weights to the 'broken leg' events or factors, and doing so inconsistently over cases (Grove & Meehl, 1996).

Hart (1998) also argues that an actuarial prediction method, calibrated and even validated on a certain sample, may not possess the same psychometric and predictive characteristics in other samples and settings. As such, wholesale adoption of an actuarial model into a setting that has distinct differences from the validation sample may be untenable. Again, Grove & Meehl (1996) counter this point. They assert that there is little evidence to support the notion that an actuarial instrument will not generalize to other settings. Using their example, they are probably right. They argue that if an instrument has been shown to work in several psychiatric clinics, there is little reason to suspect that it will not generalize to another psychiatric clinic. Even if accuracy decreases by 10%, they assert, the actuarial approach will still be superior to clinical judgment.

What if, however, settings are substantially different? Can a risk assessment tool that is validated on forensic psychiatric patients be applied with confidence to medium security prisoners? This is a researchable question, the answer to which cannot be taken for granted or assumed. In some cases, it may seem quite reasonable to decide on the basis of careful and logical thought, attention to demographics of samples, the nature of predictions being made, the consequences of those decisions, and so forth, that a certain actuarial tool can be applied at least tentatively. However, in certain cases, there may be convincing evidence that it might very well not apply, and should not be used. In response to this point, Grove and Meehl argue: either revalidate the prediction model, or develop your own.

Clearly, this debate is complex and has not subsided. The present authors' view is that, generally, actuarial prediction has been demonstrated to be superior to clinical prediction purely in terms of predictive validity. However, this statement has qualifications. Their argument is not so much that global clinical judgment would be superior to actuarial judgment when, say, the actuarial method is deemed inapplicable to a certain setting. This is not a useful comparison (i.e. global clinical impression vs. an inapplicable algorithm). Rather, it is argued here that *empirically validated, structured decisions* can be used to avoid some of the problems described above with actuarial prediction: concern over generalization to a certain setting; failure to account for salient idiosyncratic risk markers; and tendency to focus only on prediction to the exclusion of management, prevention and amelioration issues. By the term 'empirically validated structured decisions' the authors mean assessments that are guided by factors that have received empirical support in the literature. This approach is described fully below. Empirically validated, structured clinical assessment (as opposed to unstructured 'in the head', impressionistic, global clinical judgment) brings in the notion of 'conditional prediction' (Mulvey & Lidz, 1995). That is, a person may be considered to be at high risk given certain other

conditions (lack of community supervision; medication noncompliance), but at lower risk under certain conditions (access to appropriate treatment; supportive family). Reed (1997) discusses the importance of specifying which future circumstances may aggravate or mitigate risk, and whether proposed interventions can reasonably be expected to reduce risk. Further, this approach recognizes that violence risk is not static, but rather is dynamic and fluctuates with changing conditions. As such, as pointed out above, and as others have noted (Reed, 1997; Snowden, 1997), risk assessment, ideally, is a process that continues over time. Actuarial prediction typically involves the offering of a prediction index of some sort at time *A*, and then passive observation to confirm or refute the prediction at some time, *B*. Conversely, the view here is that risk assessments should reach into the future, specifying level of risk as functions of possible conditions. Treatment and management planning can then be organized specifically to disconfirm the decision of high risk.

The present authors' view, then, is that something more than mere prediction is normally needed. The use of actuarial devices, once it is decided that they are appropriate for the sample and setting, may be important *as part of the overall assessment procedure* to help inform risk level decisions. Use of such actuarial instruments, however, does not abrogate clinicians' responsibilities for their assessments and opinions. A broad approach is called for, one that goes beyond passive prediction and involves risk management, risk re-evaluation, recognition of 'critical items',[2] and statements about future events that may aggravate or mitigate risk. This approach, however, ought also to be grounded in scientifically validated knowledge. As Snowden (1997) states: '[i]f one makes important treatment or management decisions without adequate information, this is a clinical *gamble*' (p. 32; emphasis in original). The approach ought to be structured to promote systematic assessment of violence factors. These features are what separate scientifically validated structured clinical assessments from mere global clinical impressions. It is systematic, data-based, yet not so inflexible that it cannot find practical application.

Although the process of empirically validated, structured clinical assessment is discussed more fully below, the authors' general meaning is that the broad assessment approach described here ought to be rooted in knowledge that, for the most part, has been subjected to and validated by research. They do not advocate simple adoption of algorithmic prediction. As explained, their conception of risk assessment includes but goes beyond prediction. What they do insist upon, however, is the use of an assessment model that is grounded in well-substantiated research. Risk assessment schemes or guides that have been developed from a consideration of the research may improve both the clinical practice of risk assessment as well as risk assessment technology (Borum, 1996). These ideas are introduced here to provide a context for the remainder of the article. Two important questions are now discussed: (1) how accurate is risk assessment?, and (2) what predicts violence?

---

[2]These, as explained by Kropp, Hart, Webster & Eaves (1998), are items that are of such immediate salience so as to warrant, on their own, decisions of high risk. An example may be homicidal ideation.

Validity of violence prediction

Having outlined conceptual and theoretical issues in risk assessment in some detail, the validity (accuracy) of violence prediction is summarized here. The authors do not wish to suggest that this is an easy task. As Hart (1998) adeptly describes, there are many complex decisions that go into planning a single risk assessment study. To provide a 'simple' summary of validity data is not so simple. Perhaps the best starting point is where others often start—with the few studies that existed in the 1960s and 1970s.

As Cocozza & Steadman (1976) put it, in the 1960s and 1970s there was 'clear and convincing evidence' (p. 1084) for the inability of mental health professionals to predict violence accurately. In fact, Monahan's (1981) review of the topic revealed that only one in three positive predictions of violence were accurate. Since that time, given the pressure placed on mental health professionals by courts and statutes to predict violence (Douglas, Macfarlane & Webster, 1996), a good deal of research on the topic has been carried out. Conclusions from recent reviews of the subject are decidedly more positive. The increase in predictive accuracy is due at least in part to the use of improved methodologies and statistical analyses.

Studies appearing over the last few years have used better methodology than their forerunners. They have obtained correspondingly larger effect sizes. One of the important methodological advances was comprehensive follow-up procedures aimed at detecting more of the community violence that was actually occurring (see e.g. Mulvey *et al.*, 1994; Steadman *et al.*, 1998). Lines of study were published in the 1980s that showed that predictions of violence could reach at least moderate levels of accuracy.[3]

McNiel, Binder and colleagues have published numerous studies starting in the 1980s which have investigated various correlates of psychiatric inpatient violence (Binder & McNiel, 1988, 1990; McNiel & Binder, 1989, 1994*a*), and the accuracy of predictions of violence (McNiel & Binder, 1991, 1994*b*, 1995; McNiel, Binder & Greenfield, 1988). Some of these studies investigated clinical (as opposed to actuarial) prediction (McNiel & Binder, 1991, 1995; McNiel *et al.*, 1988). These studies have shown consistently that certain diagnoses, symptoms, and previous aggressive behaviours are related to inpatient violence. Also, McNiel & Binder (1994*b*) developed a screening instrument, guided by their numerous investigations, for assessing inpatient violence. It improved over chance by 25% (i.e. relative improvement over chance = 25%). Klassen & O'Connor (1988*a*,*b*,*c*, 1989) conducted several studies on the risk of community violence. They identified several domains of variables[4] which related prospectively to violence. Using this measure, 75.8% of patients were correctly classified as violent or non-violent. A moderate size multiple regression coefficient (.32) was obtained. Menzies, Webster & Sepejak (1985) developed the Dangerous Behaviour Rating Scheme (DBRS), consisting of 22 items such as anger, rage, hostility and environmental support. This instrument correlated at .34 with later violence.

---

[3]According to Cohen (1988), a moderate effect size in correlation terms is $\pm$ .30, and a large effect size is $\pm$ .50.
[4]These are: (1) early family quality; (2) current intimate relationships; (3) arrest history; (4) admissions history; (5) assault in the presenting problem (Klassen & O'Connor, 1989).

Studies appearing in the 1990s have shown sustained and even enhanced levels of predictive accuracy. The large-scale MacArthur risk assessment project (Steadman *et al.*, 1994) has produced a clinically-relevant actuarial violence prediction tool that relies upon a 'regression tree' approach to classify civil psychiatric patients into various risk categories (Monahan *et al.*, in press; Steadman *et al.*, in press). This approach requires the clinician to ask certain questions and consider certain risk markers depending on the answer to previous questions. A large number of variables (18) entered into logistic regression-based classification trees. This approach was able to obtain a strong relationship with violence.

Harris, Rice & Quinsey (1993; see also Quinsey, Harris, Rice & Cormier, 1998; Rice, 1997; Rice & Harris, 1995) found that a set of 12 variables, selected statistically from a much larger set, produced a multiple correlation with violence of .44. Coefficients in some groups were as high as .53. Lidz, Mulvey & Gardner (1993), in a well-designed study, showed that clinicians could actually offer predictions of violence that exceed chance. This group (Gardner, Lidz, Mulvey & Shaw, 1996) also developed a screening instrument for use with psychiatric patients. Using four variables (young age, multiple [3 +] previous violent acts, heavy drug use, and hostility) in a 'regression tree' design akin to that used by the MacArthur group, the researchers were able to demonstrate excellent specificity[5] (true negative rate was 99%), but poor sensitivity[6] (true positive rate was 7%). Douglas *et al.* (in press) found that the HCR-20 violence risk assessment scheme, discussed in greater detail below, was consistently and strongly associated with violence in a sample of 193 civil psychiatric patients. One promising and robust finding is the relationship between psychopathy as measured by the Hare Psychopathy Checklist—Revised (PCL–R; Hare, 1991) or Screening Version (PCL:SV; Hart, Cox & Hare, 1995). Two recent meta-analyses have shown that the link between psychopathy and violence is at least moderate (Hemphill, Hare & Wong, 1998) to large (Salekin, Rogers & Sewell, 1996).

In the 1990s, there has been a recognition of the value of statistical techniques apart from correlation, regression and classification models. For instance, Mossman (1994) has recommended the relative advantage of receiver operating characteristic (ROC) analyses over other techniques because they are much less dependent on base-rates of violence. ROC analyses produce a statistical index called the 'Area under the Curve' or AUC which is derived from plotting sensitivity against specificity for various cut-offs on a predictive measure. The AUC can range from 0 (perfect negative prediction) to .50 (chance prediction) to 1.0 (perfect positive prediction). The AUC represents the probability that a randomly chosen, actually violent person will score greater than a randomly chosen, actually non-violent person. AUC values in the mid-.70s and up may be considered large effects. Researchers have used these advantageously (Douglas *et al.*, in press; Grann, Belfrage & Tengström, in press; Rice & Harris, 1995; Strand, Belfrage, Fransson & Levander, 1999). Other techniques, such as Item Response Theory (IRT) analyses

---

[5]Specificity or true negative rate is the proportion of persons who are not violent who were not predicted to be violent.

[6]Sensitivity or true positive rate is the proportion of persons who are violent who were predicted to be so.

are useful for identifying which items in a scale carry the most information about the underlying characteristic in question (Cooke & Michie, 1997).

In his quantitative review of some violence risk assessment studies, Mossman (1994) determined that predictive accuracy was improved in more recent studies over older studies, that actuarial prediction outperformed clinical prediction, and that, overall, predictive accuracy across 58 data sets was appreciable. Current reviews, then, along with some recent risk assessment studies, suggest that violence can be predicted with moderate to large effect sizes. This statement is possible largely owing to the improved methodology used in the 1980s and 1990s. Researchers have been able to identify higher base rates of violence than in the past because of more sensitive data detection methodology (e.g. see Convit *et al.*, 1988; Mulvey & Lidz, 1994; Steadman *et al.*, 1998). Also, statistical techniques such as ROC analyses, survival curves and hierarchical regression are well-suited to identifying relationships that exist in the data.

What does 'large effect size' translate into in 'real terms', however? Some examples will help. Correlation (or regression coefficients) in the area of .50 are considered large effect sizes (Cohen, 1992). A correlation of .50 translates roughly into an odds ratio of 9.0. This would mean that a person who is predicted to be violent is nine times more likely to be violent than a person not so predicted. Douglas *et al.* (in press) found an AUC of .80 for the HCR-20 violence risk assessment scheme for predicting violent crime (Webster *et al.*, 1997*b*). This means that the probability is .80 (80%) that a randomly chosen, actually violent person will score above a given cut-off of the HCR-20, and a non-violent person will score below it. The base rate of violence in this case (violent criminal recidivism) was only 9.8%. This indicates that the prediction was improved vastly over chance. In fact, persons scoring above the median of the HCR-20 were 13 times more likely to incur violent criminal charges than those scoring below the median. Similarly, Rice (1997) reports that the AUC for the Violence Risk Appraisal Guide (VRAG) was .76, which was equivalent to a Cohen's *d* of 1.0. The MacArthur risk assessment studies (Monahan *et al.*, in press; Steadman *et al.*, in press) have found AUC values of up to .80 and .82. Again, these represent large effect sizes, and considerable accuracy in prediction. Mossman's (1994) quantitative review, which does not include the most recent and promising research, found a median AUC of .73.

To take a convincing example from outside the risk assessment arena, Cohen (1988) describes that large effect sizes are clearly discernible from observation of the data. He cites as one objective example of a large effect size the difference in height between 13- and 18-year-old women. When compared to other disciplines, how do the effects obtained in risk assessment research 'measure up'? Lipsey & Wilson (1993) carried out a meta-analysis of over 300 meta-analyses—a meta-meta-analysis. Some of the effects that they found are shown in Table 1. What seems clear is that some of the effects seem surprisingly low. For instance, chemotherapy for breast cancer produces a small effect size, as does the use of cyclosporine for organ transplants. Should, therefore, patients be dissuaded from accepting such treatments? The answer is clearly 'No' since when these effects are translated into real-life terms, they are in fact quite substantial. For instance, a correlation between treatment and recidivism of − .20 equates to a 20% drop in the chance

Table 1. Comparison of effect sizes across disciplines

Type of study	Effect (d)
Medical	
Effect of by-pass surgery on angina	.80
Use of AZT for AIDS	.47
Cyclosporine for organ transplants	.30
Chemotherapy for breast cancer	.08–.11
Drug treatment for arthritis; various outcomes	.45–.77
Psychological	
Psychotherapy, general[a]	.76
Psychotherapy, cognitive behavioural/behaviour modification[a]	.64
Treatment programmes for offenders[a]	.36
Effect of ECT on depression	.80
Drug treatment for behavioural disorders . . .	.28–.74
Drug treatment for hyperactivity; cognitive . . .	.47–.96
Neuroleptic drugs for dementia; effects on agitation	.37
Hypertensive drug therapy; effects on quality of life	.11–.28
Risk Assessment[b]	
Douglas *et al.* (in press)[c]	.91–1.07
Kropp *et al.* (1999)[d]	1.70
Rice (1997)	1.0
Strand *et al.* (1999)	1.19

Note. Adapted from Table 6 of Lipsey & Wilson (1993).
[a]Mean effect sizes calculated by the present authors from the data presented in Lipsey & Wilson (1993).
[b]Effects not from the Lipsey & Wilson (1993) meta-analysis.
[c]Calculation, though based on data by Douglas *et al.* (1998), not presented in their manuscript.
[d]Effect was calculated by the present authors from data presented in Kropp *et al.* (1998), and indicates difference in Total Scores on the SARA between groups of inmates who had or did not have spousal assault histories.

of reoffending. Over data pooled across thousands of persons, the overall benefit of treatment cannot be doubled. When the effect sizes listed in the upper parts of Table 1 are compared to the recent effect sizes being obtained in risk assessment research, it appears that this field is faring well relative to others. This is not an endorsement for complacency. Rather, it serves to demonstrate that risk assessments can indeed be made with respectable accuracy. The old question—can violence be predicted?—is not a good one. More interesting and practical is the somewhat more convoluted question: Which groups of subjects, with which particular characteristics, followed over what periods of time, are likely to exhibit precisely defined kinds of violent behaviour (see Jackson, 1997)?

Substance of risk assessment

What predicts violence? Several major factors have been shown to predict violence. Most of these have been found to cut across types of samples and settings, and are discussed globally. Although literally hundreds of specific variables have been

demonstrated to associate with violence, the following focuses on several higher-order constructs that have been shown in various studies to relate to violence with some robustness. In focusing on these groupings of risk factors, the authors are aware that others could be included, and that this discussion is not exhaustive. Readers are referred to more comprehensive accounts (Borum, 1996; Douglas & Webster, 1999a; Lyon *et al.*, in press; Monahan & Steadman, 1994a; Webster *et al.*, 1997a,b).

Historical and static factors. Under the broad category of historical and static items are included diverse risk factors. Some historical factors are simply indices of past behaviour. For instance, current young age (Harris *et al.*, 1993; Swanson, 1994), young age at time of first violence (Harris, Rice & Cormier, 1991; Harris *et al.*, 1993; Hodgins, 1983), and a history of violent behaviour (Binder & McNiel, 1990; Klassen & O'Connor, 1989; McNiel & Binder, 1994a,b) all have been shown in numerous studies, primarily of personality and mentally disordered persons, to predict violence. Maladjustment at an early age in various contexts has been shown to predict violence (Convit *et al.*, 1988; Harris *et al.*, 1993; Klassen & O'Connor, 1989).

Other historical or static risk markers may be rooted in adult past, but, like the childhood factors already mentioned, cannot be changed. A history of substance abuse is clearly a very strong marker for violence (Swanson, 1994; Steadman *et al.*, 1998). Similarly, though not without controversy and negative findings (Bonta, Law & Hanson, 1998; Harris *et al.*, 1993; Steadman *et al.*, 1998), major mental illness, generally, elevates the odds for violence (Douglas & Hart, 1999; Lindqvist & Allebeck, 1990; Swanson, 1994; Wessely, Castle, Douglas & Taylor, 1994). So too does a diagnosis of personality disorder (Douglas, Ogloff & Nicholls, 1997; Harris *et al.*, 1993) and, in particular, personality disorders characterized by anger, impulsivity and hostility (Kropp *et al.*, 1999). Examples of adult maladjustment that predict violence are relationship instability (Harris *et al.*, 1993; Klassen & O'Connor, 1988b; Saunders, 1992) and employment problems (Menzies & Webster, 1995; Shaffer, Waters & Adams, 1994). For people who already have been processed by the mental health or criminal justice systems, having attempted or completed an escape from custody, or more generally violating the terms of a conditional release, predicts violence (Harris *et al.*, 1993; Hoffman & Beck, 1985).

An important risk factor that, in theory, is relatively stable over the life-span, is psychopathy. First discussed in its modern conception by psychiatrist Hervey Cleckley (1941), the clinical construct of psychopathy has been refined, developed, and set to reliable and valid measurement by Hare (1991, 1993, 1996). Hare (1991) developed the PCL–R as a measure of psychopathy. There have been dozens of studies carried out in support of its relationship with violence and other antisocial acts. Two recent meta-analyses summarize the literature well, and, as previously noted, place the effect size between psychopathy and violence between moderate (Hemphill *et al.*, 1998) and large (Salekin *et al.*, 1996).

Dynamic and clinical factors. Although historical factors have more substantiated support behind them than dynamic, changeable aspects of persons, various clinical

constructs have also been found to predict violence. For instance, lack of impulse control and behavioural inhibition, or impulsivity—one of the items on the PCL–R, and the HCR-20—is often thought to be a strong correlate of violence (Barratt, 1994; Hollander & Stein, 1995; Webster & Jackson, 1997). In a related way, antisocial attitudes and beliefs (Andrews & Bonta, 1995), anger (Novaco, 1994) and hostility (Menzies & Webster, 1995) are states that, if currently afflicting a person, will probably elevate violence risk. Whether a person is amenable to treatment, or rather is resistant to remediation attempts, also has relevance for violence (Bartels, Drake, Wallach & Freeman, 1991).

Although, as discussed earlier, some researchers have found that mental illness, on a global level, does not relate to violence, or in fact relates negatively, numerous studies have found that persons with major mental disorders are actually at an increased risk for violence. Some have suggested that it is the active, florid states of mental illness that are the key ingredient in the relationship between violence and mental illness (Buchanan, 1997; Douglas & Hart, 1999; Monahan, 1992). For instance, certain aspects of delusions, such as being able to identify evidence or support for the belief, and being made afraid because of the belief, have been shown to elevate the chances that a person will act on a delusion (Buchanan, 1997). A concept known as threat-control override (TCO) psychotic symptoms, which involves the perceived loss of control over decisional functions (e.g. thought insertion) in concert with perceived threats to personal safety (e.g. beliefs that others are following and out to cause one harm), has been shown to be related to violence (Link & Stueve, 1994; Swanson, Borum, Swartz & Monahan, 1996). In a meta-analysis of over 100 studies on the topic of psychosis and violence, Douglas & Hart (1999) found that not only was there an association between major mental disorders and violence generally, but that this relationship was stronger when analyses used symptoms as the level of measurement, rather than the more gross categorization of diagnosis of disorder. Other dynamic markers for violence risk include suicidal or self-harm intent (Hillbrand, 1995), sadistic fantasies (MacCulloch, Snowden, Wood & Mills, 1983) and homicidal ideation.

Risk management factors. This category of violence risk markers has substantial importance to the violence risk assessment task in that these variables, if targeted, can help to ameliorate risk. Since they are, in fact, changeable, these markers could be included under the previous section. However, their distinguishing qualities are that they tend not to relate to characteristics of persons themselves, but rather to aspects of the situation or environment in which a person will be living upon release (or at least at some future time, or to the person–situation interaction).

For instance, whether a release plan is feasible (i.e. developed to suit a person's particular requirements) may have a bearing on violence risk (Andrews, Zinger, Hoge, Bonta, Gendreau & Cullen, 1990). Unfeasible, poorly thought out or weak plans are likely to increase violence risk. Further, people who have access to the appropriate amount and type of professional help are at reduced risk to act violently (Estroff & Zimmer, 1994). In a parallel manner, personal support from family and friends is an important part of ameliorating violence risk (Estroff &

Zimmer, 1994; Klassen & O'Connor, 1989), so long as family and friends are understanding of relevant mental health and violence issues. Obviously, such support must act as buffers to violence relapse, rather than as triggers to ignite it. Concerning triggers in a different light, it is important that assessors evaluate whatever destabilizers are likely to aggravate violence risk. Such factors are idiosyncratic and personalized depending on a person's criminogenic propensities. For example, a paedophile who is released to live beside a school will encounter, on a daily basis, children playing and coming to and from school. This circumstance is apt to be destabilizing, and surely increases violence risk. Finally, salient stressors may make coping with the demands of everyday life less manageable than otherwise, and hence increase the probability of poorly judged or thought out reactions, including possibly violent ones (Felson, 1992; Klassen & O'Connor, 1994).

Process of risk assessment: empirically validated structured clinical decision-making

Much of the above discussion of risk assessment focuses on the science of the task. How does this translate into clinical action? As should be clear by this stage, an ordered, systematic approach to structured clinical decision-making is endorsed. This may involve first deciding if any or all of a number of actuarial schemes are suited to the task. Such an analysis requires thoughtful comparison between the characteristics of the individual under assessment and the features of the population for which the actuarial scale was developed. It may even be necessary, locally, to array data from comparable individuals and determine the extent to which the individual under consideration does or does not fit group-based characteristics. If an actuarial scheme is considered applicable, it should be used *as part of the overall assessment procedure*. It may be an important data point for a clinician to consider as he or she forms an opinion about violence risk.

   Use of actuarial schemes does not remove the necessity of clinical decision-making. There most likely will remain the legal and ethical obligation to ameliorate violence risk after making predictions of 'high' violence risk (Hart, 1998). That is, decisions regarding risk management and prevention are required, and are as crucial as the first step of making a violence prediction. The use of actuarial tools also does not remove the necessity of clinical skill and judgment. For instance, the VRAG (Harris *et al.*, 1993), an actuarial scare that has been shown to relate strongly to violence, requires ratings for psychopathy, schizophrenia and personality disorder. Diagnosis of these conditions is an inherently clinical task. Clinicians also have to decide when an actuarial process is not indicated. The VRAG was calibrated on a sample of male forensic psychiatric patients. Whether its predictive properties generalize to, say, female prisoners or to other settings is a researchable but as yet unanswered question. Its use in certain settings is ruled out by its own construction, in that it contains several items that presume the existence of an index offence (e.g. offence severity, gender of victim). In some settings where violence is a clinical and legal concern though is not inherent to the setting (e.g. civil psychiatric), these variables may not apply, as persons may be civilly committed without having perpetrated actual violence against others.

The use of an empirically based guide or *aide-mémoire* to structure the clinical assessment is forwarded as an appropriate and comprehensive violence risk assessment model. Several of these that have, of late, been developed rationally from consideration of the scientific and professional literatures are reviewed here. These instruments, though capable of being used as actuarial predictors, are better conceptualized as assessment protocols. These clinical guides embody what is meant by 'empirically validated structured clinical decision-making'.

The rationale underlying such instruments is to integrate the tasks of prediction, assessment, management, prevention and communication. Four schemes are highlighted that may be described as guides for structured clinical decision-making: the HCR-20; the Spousal Assault Risk Assessment (SARA) guide; the Sexual Violence Risk-20 (SVR-20); and the Early Assessment Risk List for Boys (EARL-20B; Augimeri, Webster, Koegl & Levene, 1998). The first of these, the HCR-20, is meant to apply to assessments for violence generally. The SARA was designed for the assessment of spousal violence risk. The SVR-20 was created for assessing risk for sexual violence. The EARL-20B, as its name implies, is intended to assist clinicians to assess violence risk in childhood.

These schemes have commonalities. At a conceptual level, they are manifestations of the scientist–practitioner philosophy, in that each possesses risk factors that were selected through review of the scientific and professional literatures, and through consultation with clinicians in respective fields. They were conceived rationally, or analytically. Each was designed to be used by clinicians, as well as to be testable according to conventional research standards. Items can be scored by clinicians in an applied setting, and may also be coded reliably by researchers (or researcher-clinicians) in research applications. Making these schemes 'user friendly' was a driving concern in their development. Complex scoring procedures and cumbersome assessment protocols cannot feasibly be integrated into routine clinical practices. The four risk assessment guides discussed here were created to be easily adopted into day-to-day clinical regimes. Some adaptation to suit local circumstances was anticipated. Although all were devised as guides or *aide-mémoires*, each also is capable of being treated as an actuarial device, in that numbers can be aggregated in some fashion and linked to subsequent violence. However, these risk assessment schemes are intended to include, but be broader than, instruments designed principally to offer a numerical index of risk.

Each instrument recommends a certain method of use. Generally, persons are rated on the presence of the relevant risk factors. Then, a 'summary risk rating' of high, moderate or low risk is given, based on the clinician's overall appraisal. This approach clearly does call for clinical judgment—but it is structured. Advocates of pure actuarial approaches would probably label this approach as 'clinical judgment'. The present authors agree. Yet they assert that there may be defining differences between clinical opinion, as conceptualized by Meehl (1954) and Grove & Meehl (1996) and the present approach to structuring assessment procedures. Grove and Meehl call clinical prediction impressionistic and subjective. The present authors think that the approach propounded in this paper need be neither impressionistic nor subjective. It also is not necessarily algorithmic or mechanical, as actuarial prediction is defined to be. Recent research supports the utility of this approach.

Using the SARA, Kropp *et al.* (1999) showed that the 'summary risk ratings' made by clinicians (which would fall into the 'clinical' camp in the traditional debate) actually outperformed the relationship between the sum of the instrument's individual items and violence (actuarial prediction). As Kropp *et al.* note, this was the first time, to their knowledge, in the violence risk assessment field, that 'clinical' judgment outperformed actuarial methods. Dempster (1998) also found that structured clinical ratings of risk using the SVR-20 added incremental validity to actuarial predictions. While these findings do not, of course, overturn the many findings that have found actuarial prediction to out-predict clinical impressions, they do provide some support for the structured clinical decision-making approach.

In this family of violence risk assessment schemes that follow, clinicians are called upon systematically to assess violence risk markers, indicate their presence, and then assign an overall risk rating based on the importance of these factors for the person being assessed. These tools also feature the option of 'critical items', which are variables that are of such salient importance for the current assessee that, alone, they warrant ratings of high risk. Clinicians are encouraged to base risk ratings on the items, but there is no specific concrete recommendation for cut-offs. Part of the reason for this stems from the fact that a person can be at high risk with the presence of a single (critical) risk marker. However, clinicians may devise their own procedures for rating people as high, moderate and low risk that are guided by score ranges, number of items present, and types of items present. One of the schemes, the SARA (Kropp *et al.*, 1999), provides guidance through the display of percentile tables and by placing certain scores into 'zones' of risk.

It is the present authors' contention that the flexibility afforded by this approach to risk assessment is an advantage when it comes to dealing with treatment and rehabilitation issues. Though flexible, the assessment method also is structured, guided and empirically based. As discussed earlier, this approach recognizes that the clinician is ultimately responsible for decisions, that (existing) actuarial tools may not be entirely appropriate and completely relevant in some settings, and that actuarial tools may fail to allow clinicians to 'flag' important markers in individual cases. If an individual issues definite and serious threats to kill some other person, but 'homicidal ideation or intent' did not make it into the regression equation, then the individual could probably erroneously be rated as low risk.

Violence risk assessment schemes

HCR-20. The HCR-20 applies to the assessment of violence risk generally (i.e. it does not target but includes sexual violence and spousal violence). Its three scales, as shown in Table 2, capture past (Historical), present (Clinical) and future (Risk Management) aspects of violence risk. Each of the 20 items are scored 0, 1, 2. A score of 0 indicates that the item is definitely absent; a score of 1 indicates that the item is possibly present or present in less serious form; a score of 2 indicates that the item is definitely present or present in more serious form. This scoring procedure was adopted from the PCL–R (Hare, 1991) because it appears to promote reliability, and it is easily understandable by clinicians and researchers

Table 2. Items in the HCR-20 risk assessment scheme

Subscales	Items
Historical scale	
H1	Previous violence
H2	Young age at first violent incident
H3	Relationship instability
H4	Employment problems
H5	Substance use problems
H6	Major mental illness
H7	Psychopathy
H8	Early maladjustment
H9	Personality disorder
H10	Prior supervision failure
Clinical scale	
C1	Lack of insight
C2	Negative attitudes
C3	Active symptoms of major mental illness
C4	Impulsivity
C5	Unresponsive to treatment
Risk management scale	
R1	Plans lack feasibility
R2	Exposure to destabilizers
R3	Lack of personal support
R4	Noncompliance with remediation attempts
R5	Stress

Note. Adapted from Webster, Douglas, Eaves & Hart (1997*b*).

alike. Each of the 20 items enjoys at least some backing from current published research. Support for the validity of the HCR-20 has been found in samples of civil psychiatric patients, forensic psychiatric patients, and correctional inmates. Much of this research is unpublished, given the HCR-20's relative youth.

There have been two civil psychiatric samples that have yielded several manuscripts and presentations. Douglas *et al.* (in press) completed the HCR-20 on 193 psychiatric patients who were released into the community. Using ROC analyses, AUCs ranged from .76 (for any violent act and for physical violence) to .77 (for fear-inducing or threatening behaviour), to .80 (for arrest for violent offences). These effects are large in size. In fact, persons who scored above the median of the HCR-20 were 13 times more likely to be arrested for violent offences than those who scored below the median. The PCL:SV (Hart *et al.*, 1995) also was coded in this study, and it too produced respectable AUCs (range from .68 to .79), as well as an odds ratio of approximately 13 for violent offences. A direct comparison of the two instruments revealed that the HCR-20 produced a statistically significantly greater AUC for the prediction of any violence and for fear-inducing/threatening behaviour. Regression analyses revealed that the HCR-20

and its scales were more consistently related to violence than the PCL:SV. Nicholls, Ogloff & Douglas (1997) found that, generally, the HCR-20 performed well for women in addition to men, although it was not as strongly related to the physical violence of women (AUC = .63) as it was for men.

In another civil psychiatric setting, Ross, Hart & Webster (1998) found, in their sample of 131 acute psychiatric inpatients, that AUCs ranged from .68 to .75 for the prediction of community violence. For inpatient violence in this sample, AUCs were smaller (.63 to .68), although still of moderate size and significantly greater than chance. The PCL:SV also was coded for this study, and its AUCs were somewhat smaller than those for the HCR-20, ranging from .65 to .70 for community violence, and averaging .61 for inpatient violence (although the statistical significance of differences was not reported).

Reliability in these civil psychiatric settings is acceptable. Douglas *et al.* (in press) report an overall intraclass correlation coefficient (ICC_1) between raters of .80. Ross *et al.* (1999) also found inter-rater reliability to be .82. In terms of internal reliability or consistency, Cronbach's alpha was reported by Ross *et al.* to be .74 for the Historical scale, .64 for the Clinical scale, and .69 for the Risk Management scale.

In forensic psychiatric samples, seven data sets have provided various disseminations on reliability and validity. Five of these samples have offered some validity data—for instance, Strand, Belfrage, Fransson & Levander (1999), in a Swedish sample of 40 male forensic psychiatric patients,[7] 22 of whom were recidivists, and 18 of whom were non-recidivist, matched controls. This research also used the PCL:SV. The AUC for the HCR-20 was .80, and for the PCL:SV, .70. The mean HCR-20 score for the recidivist group was 30.8, and for the non-recidivist group 22.4. This eight-point difference translates into a Cohen's *d* of 1.19, which is clearly a large effect size.[8] Every person who scored above 34/40 on the HCR-20 ($N = 11/40$) recidivated. Surprisingly, items from the Clinical and Risk Management scales were most powerful in terms of discriminating the two groups. This may stem from the fact that, as Strand *et al.* report, the two groups were quite homogeneous on the Historical items.

In another Swedish sample, Grann, Belfrage & Tengström (in press) compared the Historical scale of the HCR-20 to the VRAG (Harris *et al.*, 1993) in 404 forensic patients and mentally disordered offenders. They carried out analyses on the entire sample, as well as on subgroups of schizophrenic patients ($N = 293$) and personality disordered patients ($N = 111$). The Historical scale AUC for the entire sample was .71, and for the VRAG it was .68. Similarly, in the personality disordered group, the AUC for the Historical scale was .71, and for the VRAG it was .68. In the schizophrenic cohort, the AUC was lower for the Historical scale, at .66. The AUC for the VRAG was not greater than chance.

---

[7] This research project, in addition to all others carried out in Sweden, used the Swedish translation of the HCR-20 (Belfrage & Fransson, 1997). This is a literal translation of Version 2 of the HCR-20 (Webster *et al.*, 1997*b*), and hence findings based upon it ought to be interpretable as support for the HCR-20 generally.

[8] This *d* was calculated by the present authors. According to Cohen (1988, 1992), a *d* score (which is $[M_1 - M_2]$/pooled SD) of .80 + is considered large in magnitude. A *d* of 1.19, as in Strand *et al.*'s study, means that the difference between the groups' means is 1.19 times their average standard deviation.

In a Canadian sample of 80 male forensic psychiatric patients, Wintrup (1996; see also Douglas, Webster & Wintrup, 1996) found that both the HCR-20 and PCL–R correlated at approximately .30 with several measures of later community violence. The HCR-20 was quite strongly related to subsequent re-admissions to the forensic facility ($r = .38$) and to civil psychiatric hospitalization ($r = .45$).[9]

Finally, in another Canadian sample of 175 consecutively assessed forensic psychiatric patients, Douglas, Klassen, Ross, Hart & Webster (1998) studied the concurrent validity of the HCR-20 through its relationship with the PCL–R, the Brief Psychiatric Rating Scale (BPRS; Overall & Klett, 1962), and indices of violent and other antisocial behaviour. The HCR-20 total score correlated strongly with both the PCL–R ($r = .61$) and BPRS ($r = .54$). However, the relationship between particular HCR-20 scales and these other instruments differed importantly from the overall correlation. For instance, the Historical scale was very strongly related to the PCL–R ($r = .75$) whereas the Clinical and Risk Management scales were related with small/moderate effects (rs = .21 and .18, respectively). Conversely, the latter two scales were strongly related to most of the factors on the BPRS (rs = .63 and .59), whereas the former was not. This pattern of findings makes conceptual sense, given the content of the various scales. With respect to the HCR-20's relationship with criminal record, people scoring above the median of the HCR-20 were approximately two to four times more likely to have criminal histories of violence offences, past assault charges, and juvenile criminal histories than those scoring below the median. For these analyses, Item H1 (Previous violence) was removed from analyses to prevent artificial inflation of the effects.

Concerning the reliability of the HCR-20 within forensic samples, Belfrage (1998), in another Swedish sample, had six clinicians rate the same 43 patients. Using multivariate inter-relater reliability analyses, Kendall's $W = .81$ for the total scale. In this same study, internal consistency for the total scale, using Cronbach's alpha, was .95. In the Douglas *et al.* (1998) sample, inter-rater reliability was determined using the ICC, both single measure (ICC_1) and average measure (ICC_2). Here, $ICC_1 = .80$ and $ICC_2 = .89$. Also in this study, Cronbach's alpha for the total scale was .79. In a sample of 100 forensic patients in Germany, Müller-Isberner & Jöckel (1997)[10] found that the mean Kappa for the Historic scale was .89, with 91% of cases being rated by two clinicians within 1 point. For the Clinical scale, mean Kappa was lower, at .49, with 71% of cases being within 1 point.

Lastly, there have been three completed studies of the HCR-20 in correctional or prison samples. Belfrage, Fransson & Strand (1999) carried out a prospective study of violence within two maximum security prisons in Sweden, using the HCR-20. Eight of the 41 (19.5%) inmates were violent in the prison. The Clinical scale, Risk Management scale, and HCR-20 Total Score differentiated between the violent and non-violent groups. The HCR-20 Total Score was 33.4 in the violent group, and 24.6 in the non-violent group. All HCR-20 Risk Management scale items were

[9]According to Cohen (1988, 1992) a Pearson correlation coefficient of ± .30 is considered moderate in magnitude, and ± .50 is considered large.
[10] The German version of the HCR-20 contains three items not in the original HCR-20, and is based on Version 1 (Webster *et al.*, 1995), not Version 2 (Webster *et al.*, 1997*b*) of the HCR-20. It was not possible to rate the Risk Management scale.

significantly greater among the violent group than the non-violent group. The Historical scale was not predictive of violence, except for Item H10. In a sub-group of 30 psychopaths, defined with the PCL:SV, the Risk Management scale and HCR-20 Total Score were significantly higher in the violent inmates. Four of the five Risk Management scale items were higher in the violent psychopaths compared to the non-violent psychopaths. The authors comment that the Historical scale was not predictive in this sample because inmates (being maximum security violent inmates) were homogeneous with respect to most historical factors. The Clinical and Risk Management factors did, however, provide for a means of separating violent from non-violent inmates. These results are consistent with those of Strand *et al.* reported above. The results of the study, though limited by a small sample, provide support for the importance of risk management concerns for high-risk violent offenders.

In a second study, Douglas & Webster (1999*b*) evaluated the concurrent validity of the HCR-20 in 72 male prisoners who had been referred to a regional health centre of the federal Canadian prison system. They were rated on the Historical and Clinical scales of the HCR-20,[11] the PCL–R and the VRAG. Correlations between these scales and past violence, again with Item H1 (Previous violence) removed from analyses, were, respectively, .50 and .30 (for the HC composite, $r = .44$). The VRAG correlated at .20 and the PCL–R at .41, with past violence. Scores above the median of the HC composite increased the odds of various types of past violent and other antisocial behaviour by an average of four times. Hierarchical regression analyses showed that the HCR-20 (particularly the Historical scale) added incremental validity to the relationship between the PCL–R and violence, and the VRAG and violence. The converse was not true—that is, neither the VRAG nor the PCL–R added appreciably to the relationship between the HCR-20 and violence. This was particularly so for the VRAG. In this study, inter-rater reliability between two raters, using Pearson *r*, was .80. In another postdictive study, Dunbar (1999) observed correlations of .33 to .63 between the HCR-20 and violence in a sample of American hate crime perpetrators.

In summary, across civil psychiatric, forensic psychiatric and correctional samples, the HCR-20 has demonstrated effect sizes with violence that typically are either large or moderate in size by conventional definitions. In the few studies that have compared its validity to other measures (PCL:SV; PCL–R; VRAG) it has performed well. Reliability has ranged from acceptable to good, with most inter-rater indices being in the .80s. Internal consistency has been rated in the high .70s to the .90s.

SARA. The SARA possesses conceptual underpinnings similar to the HCR-20, but is intended to be applied in cases where there is a known or suspected background or concern about violence against spouses or partners. The SARA also possesses 20 items that are scored 0, 1, 2. These items are grouped differently from the HCR-20 items. As shown in Table 3, the SARA is divided into four types of risk factors: Criminal History; Psychosocial Adjustment; Spousal Assault History; and Alleged

---

[11]It was not possible to complete the Risk Management scale because most prisoners had not yet been considered for release from prison.

Table 3. Items in the SARA risk assessment scheme

Subscales	Items
Criminal History	
1	Past assault of family members
2	Past assault of strangers or acquaintances
3	Past violation of conditional release of community supervision
Psychosocial Adjustment	
4	Recent relationship problems
5	Recent employment problems
6	Victim of and/or witness to family violence as a child or adolescent
7	Recent substance use/dependence
8	Recent suicidal or homicidal ideation/intent
9	Recent psychotic and/or manic symptoms
10	Personality disorder with anger, impulsivity, or behavioural instability
Spousal Assault History	
11	Past physical assault
12	Past sexual assault/sexual jealousy
13	Past use of weapons and/or credible threats of death
14	Recent escalation in frequency or severity of assault
15	Past violation of 'no-contact' orders
16	Extreme minimization or denial of spousal assault history
17	Attitudes that support or condone spousal assault
Alleged (Current) Offence	
18	Severe and/or sexual assault
19	Use of weapons and/or credible threats of death
20	Violation of 'no-contact' orders

Note. Adapted from Kropp, Hart, Webster & Eaves (1999).

(Current) Offence. The Criminal History section deals with past crimes generally. Psychosocial Adjustment contains risk markers that span various areas from relationships to psychosis. The Spousal Assault History section contains items that are related specifically to previous spouse assault. Finally, the section dealing with the Alleged (Current) Offence has raters consider three items that are similar to the previous spouse assault items but apply only to the present index offence.

The SARA manual (Kropp *et al.*, 1999) provides appreciable validity and reliability data. These analyses were carried out on samples totalling more than 2300 probationers and inmates. Validity was estimated according to known-groups and concurrent validity strategies. Impressively, every item, scale score and summary score had higher scores for participants with known histories of spousal assault than for those without.

In terms of concurrent validity, the PCL:SV, VRAG, and the General Statistical Information on Recidivism (GSIR; Nuffield, 1982) were correlated with the SARA. All SARA ratings (summary risk ratings, total scores, scale scores, and number of items present) were moderately to highly correlated with the PCL:SV. The GSIR

and VRAG, on the other hand, were moderately to highly correlated with only the general violence items of the SARA, but were not related to the spousal violence section scores or to summary risk ratings. Kropp *et al.* interpret this pattern of findings as support for the convergent validity of the general violence part of the SARA and for the discriminant validity of the spousal assault part of the SARA.

Finally, in a post-diction analysis of a subset of known recidivist and non-recidivist spouse assaulters, various SARA indices discriminated between recidivists and non-recidivists. For instance, recidivists had higher scores on the number of critical items present, number of items present, and summary risk ratings. Generally, differences were larger for the spousal assault section rather than the general violence section. It is important to point out that these groups did not differ on important demographic variables (e.g. age, education, employment, marital status, ethnicity), history of assault or on treatment variables (except one).

A very important finding was that summary risk ratings ('clinical' prediction) outperformed total scores ('actuarial' prediction) in terms of their relationship to spousal assault recidivism. The type of clinical prediction facilitated by the SARA is, of course, empirically guided structured decision-making, rather than the unstructured, global, impressionistic, clinical decision-making that is typically considered in debates over clinical vs. actuarial prediction.

In terms of reliability, internal consistency for the total scale, as measured by Cronbach's alpha, was .78. Item homogeneity, as estimated by mean inter-item correlations (MIC), was .15 for the total scale. Inter-rater reliability indices were very good—$ICC_1 = .84$ and $ICC_2 = .91$. Inter-rater reliability was (expectedly) somewhat lower for summary risk ratings, though still good—$ICC_1 = .77$ and $ICC_2 = .77$.

The SVR-20. As with the HCR-20 and SARA, the SVR-20 was constructed rationally from a consideration of the scientific and professional literatures on sexual violence. The obvious difference is that its purpose is the assessment of sexual violence risk. It, too, has 20 items that are scored on the 0, 1, 2 basis. A slight deviation from the HCR-20 and SARA item coding is that SVR-20 items are rated Present (Y), Absent (N) or Unsure (?). In practice and in research, however, the scoring system is the same as the PCL-R, HCR-20 and SARA. One useful addition to the scoring of the SVR-20 is the allotment of a 'recent change' score for each item. Assessors can indicate whether each risk factor has become more or less salient in some recent specified period of time. As with the other guides, the SVR-20 provides assessment guidelines, recommendations for the communication of assessment findings, and management strategies based on ratings of high, moderate or low risk. Table 4 shows that the 20 items are divided into three categories: Psychosocial Adjustment; Sexual Offences; and Future Plans.

The one completed study on the SVR-20 was carried out by Dempster (1998), who coded the SVR-20, along with other sexual violence and forensic assessment instruments, on 95 sexual offenders who had been incarcerated in federal Canadian institutions. Both the 'actuarial' (summed total of items) and 'structured clinical' (ratings of high, medium and low risk) forms of completing the SVR-20 had moderate-sized univariate correlations with sexual violence (the former correlating

Table 4. Items in the SVR-20 risk assessment scheme

Subscales	Items
Psychosocial Adjustment	
1	Sexual deviation
2	Victim of child abuse
3	Psychopathy
4	Major mental illness
5	Substance use problems
6	Suicidal/homicidal ideation
7	Relationship problems
8	Employment problems
9	Past nonsexual violent offences
10	Past nonviolent offences
11	Past supervision failure
Sexual Offences	
12	High density sex offences
13	Multiple sex offence types
14	Physical harm to victim(s) in sex offences
15	Uses weapons or threats of death in sex offences
16	Escalation in frequency or severity of sex offences
17	Extreme minimization or denial of sex offences
18	Attitudes that support or condone sex offences
Future Plans	
19	Lacks realistic plans
20	Negative attitudes toward intervention

Note. Adapted from Boer, Hart, Kropp & Webster (1998).

at .31 and the latter at .34, with sexual violence). The SVR-20 clinical ratings produced an AUC of .77 with sexual recidivism, which was equal to or greater than concurrent measures of sexual violence risk. The SVR-20 clinical ratings were one of two measures (of a total of six) that were able to distinguish between sexual recidivism vs. non-sexual but otherwise violent recidivism. A powerful finding by Dempster was that, for the prediction of sexual recidivism, the SVR-20 structured clinical ratings had incremental validity above and beyond that of various actuarial predictive scales of sexual violence risk. That is to say, using a hierarchical Cox regression model, the SVR-20 clinical ratings produced significant model improvements above the sole inclusion of the actuarial predictors. For some of these analyses, both the actuarial and clinical predictions remained significant once both predictors were entered, and for one analysis only the clinical ratings retained significance.

EARL-20B. This device evolved from the Earlscourt Child and Family Centre (ECFC) in Toronto, Canada. The agency assists children under age 12 and their families. Most of the children are referred by the police because of conduct problems. Although theft and fire-setting are often referral issues, the most

Table 5. Items in the Early Risk List for Boys (EARL-20B)

Subscales	Items
Family	
F1	Household circumstances
F2	Caregiver continuity
F3	Supports
F4	Stressors
F5	Parenting style
F6	Antisocial values and conduct
Child	
C1	Developmental problems
C2	Onset of behavioural difficulties
C3	Trauma
C4	Impulsivity
C5	Likeability
C6	Peer socialization
C7	School functioning
C8	Structured community attitudes
C9	Police contact
C10	Antisocial attitudes
C11	Antisocial behaviour
C12	Coping ability
Amenability	
A1	Family responsivity
A2	Child treatability

Note. Adapted from Augimeri, Webster, Koegl & Levene (1998).

pressing concerns usually centre around aggression and violence. The EARL-20B was developed from a longer 56-item checklist (Augimeri & Levene, 1997) and, using some of the ideas derived from the HCR-20, the SARA and SVR-20 noted above, was reduced to an assessment manual with 20 items. As with the other devices, much time was devoted to both scouring the literature and consulting with clinicians. Some colleagues will ask, perfectly reasonably, whether such a focus on the assessment of aggression in boys might be wise in persons so young, that there can be dangers in 'labelling' children so early. Issues such as these do, of course, require thoughtful and respectful consideration. But in the end it is necessary to decide whether the benefits to clients are likely to outweigh the costs and, if so, to move forward cautiously.

As can be seen in Table 5, the EARL-20B is comprised of three sections: Family Factors; Child Factors; and Amenability Factors. Family factors attempt to tap whether there is nurturance, support and supervision in the home. Child factors expand upon core constructs of sociability and spontaneity. Amenability factors are concerned with the resilience and competence of the child and family. Many of these items are similar to those found in the HCR-20 (e.g. stressors, antisocial

values and attitudes, supports). Of course, there is strong focus on issues of childhood. Many of these risk factors are considered under a single HCR-20 risk factor (H8: Early maladjustment). In childhood, however, these factors are not necessarily historical, as in the HCR-20, but may be currently affecting a child.

As with the other devices, the EARL-20B is intended to inform treatment and management decisions. The identification of critical items (as in other of the devices) does a great deal to galvanize prompt and effective interventions. It helps to sort out not just how many factors are important, but assists in settling the order in which problems might best be tackled. There is also the chance that the EARL-20B will help in determining, perhaps from total scores or numbers of critical items, which children and families must receive immediate intensive service. It may even be helpful to complete the EARL-20B with the child's parent as a way of reaching agreement with parents over the exact nature of the problem and the best way of tackling it. It should be pointed out that similar scheme for use with girls is being developed. It is likely that the sorts of issues that are important for aggression in boys and girls warrant separate attention.

Communicating the results of risk assessments

Above it was explained that violence risk assessment may be able to be carried out with some respectable degree of accuracy, and that there are schemes available to promote this. When this is so, how should a clinician communicate the findings of his or her risk assessment? What degree of precision is warranted by the empirical research? This is a major issue in the risk assessment field (see Monahan & Steadman, 1996; Schropp, 1996). In the past, predictions were often given in Yes/No terms. An individual was deemed 'dangerous' or 'not dangerous'. Such an approach is probably too simple and begs for errors. Some researchers advocate communicating violence risk estimates in terms of precise probability levels, such as .82 over 10 years (Quinsey *et al.*, 1998). Others prefer more general statements that speak of such terms as 'low', 'moderate' and 'high' risk (Monahan & Steadman, 1996; Webster *et al.*, 1997*b*).

The authors favour the latter approach—use of broad categories such as 'low', 'moderate', 'high', and 'imminent', somewhat along the lines suggested by Monahan & Steadman (1996). The present authors' argument is that the specification of precise probabilities probably conveys a level of exactitude that can be misleading. Consider the idea of 'conditional prediction' offered by Mulvey & Lidz (1995). These commentators argued that risk for violence is conditional upon a great many factors that may occur post-release, after an initial risk assessment has been completed. Risk assessment, as the present authors consider it, is by no means solely a static process. The dynamic and conditional nature of risk assessment as a continuous, changing construct belies the possibility of giving precise, static estimates of violence risk. For instance, does the probability remain precisely fixed for every person with a given score on an instrument? 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* risk for violence (see Monahan & 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 character-istics from similar institutions.

The idea of risk management discussed at points throughout this article integrates some of these ideas. For instance, a person may be considered to be at moderate risk for violence relative to, say, other persons released from the same institution, based on an assessment of historical, clinical and risk management factors. This risk may be conditional upon the risk management factors remaining the same. If, for instance, the person fails to comply with medication, absconds from treatment, and begins abusing substances, then risk may be considered to be high. The advantages of such an approach are the implications for management that flow from the assessment (Boer *et al.*, 1998; Monahan & Steadman, 1996; Webster *et al.*, 1997*b*) and the flexibility to change violence risk communications should conditions change.

It should be evident from the above that the authors do not disagree with calculating the percentage of people who act violently as a function of being in one or other of various risk categories. In fact, this exercise would be useful in order to validate the risk communication approach; it also yields descriptive information to users about risk assessment models that employ such communication statements. But, as well as this, the authors favour explicating the violence risk factors that are relevant to the case at hand, and specifying those that are changeable.

Risk assessment training

Much of what has been said above has strong implications for training. Pro-grammes that provide clinical training ought surely to include a component on violence risk, and should offer training on validated assessment methods. There is also a need to guide education about research on violence risk assessment. An example of a training module that could be found as part of a graduate, medical or other professional programme that included emphasis on forensic issues is outlined in Table 6. Introduction to both the science and practice of risk assessment proceeds from very introductory-level activities such as reading and clinical case or video review to carrying out independent research studies and handling a clinical caseload. Borum (1996) has suggested as well that violence risk assessment be declared a proficiency area within the mental health professions. It is, as he reminds us, disconcerting that training students and professionals in violence risk assess-ment is not more entrenched, given that many, if not most, professionals will come across a situation where risk assessment is called for in some manner.

Implications for clinical practice and research

What lessons for clinical practice and research follow from the ideas forwarded in this paper? In terms of practice, the implications seem relatively straightforward.

Table 6. Violence risk assessment training curriculum

Research	Practice
Level 1	Level 1
Lectures; reading	Observation, reading
Level 2	Level 2
Literature reviews	Videotape case reviews
Research proposals	Mock risk assessments/reports
Level 3	Level 3
Research assistantship	Clinical assistantship
Run pre-planned studies	Assist case preparation/file review
Assist in data processing/analyses	Assist in risk determination
Assist in write-up and dissemination	Assist in case write-up
Level 4	Level 4
Supervised minor research project	Supervised clinical risk assessment
Plan; conduct; write; disseminate	Assess; interview; write; opine
Mock/actual research presentation	Case conference/grand rounds
Level 5	Level 5
Major research project/thesis	Clinical caseload/practicum

Clinicians should base decisions on substantiated assessment procedures. This also means that they should contribute to the development and testing of such devices. Researchers should be at pains to help consolidate the scientific literature in ways that have practical use to clinicians. This means that they should have, where possible, familiarity with actual individual assessment cases and be challenged to deal with them in the light of prediction and assessment schemes.

The general framework for this article is the scientist–practitioner model. This approach, generally, holds that the practice of psychology (and by extension the other mental health professions) should be informed by scientific research. It is argued that this model has an apt fit in the risk assessment field. Despite the fact that much research knowledge is of recent origin, the authors agree with Borum (1996) that there has been enough advance of late to allow general consensus about key areas of risk assessment. Guidelines can, it would seem, be written responsibly. The authors assert further that science *ought* to inform practice. As indicated above, it may be ethically questionable for practitioners to fail to avail themselves of easily available research and technology. In tort law, professionals may be held to be negligent if they do not perform as a reasonable professional would, and injury ensues as a result. It certainly is reasonable for practitioners to be expected to be familiar with the state of their discipline in terms of practice *and* research. Ethical guidelines of professional associations require as much (APA, 1992). Failure to be so familiar could result in professional negligence. The authors do not advocate blind adoption of any research findings, but critical evaluation of research and adoption of reliable and valid findings that could be reasonably expected to generalize to the setting at hand.

The approach outlined above with the HCR-20, SARA, SVR-20 and EARL-20B does seem to be a slightly new addition to that aspect of forensic mental health practice which is concerned with assessing and managing risk for violence. This is not to say that others have not been at pains to produce guides in the past. Monahan's (1981) text was turning in that direction, and certainly the MacArthur Risk Assessment Project in the United States (Monahan & Steadman, 1994*b*; Monahan *et al.*, in press; Steadman *et al.*, 1994, 1998, in press) offers much in this regard. As a result of the present authors' experience to date, they think that this kind of conceptualization opens the way for different kinds of scientific and professional literatures to be connected to the literature on risk assessment. What they are mainly discussing here is the very large body of literature which deals with the scientist vs. practitioner debate, discussed earlier.

In advancing this alternative, somewhat more flexible and clinically collaborative line, the authors do not for a moment argue against the continuing effort to establish in more 'linear' fashion the exact links between clearly defined predictions and clearly defined outcomes. Once the basis for the four scales described here reaches a sufficient level of firmness, it seems that colleagues begin to use standard methodology to confirm the guides' reliability and validity. The authors have observed that colleagues tend to go over these instruments with painstaking care. Well in advance of wanting to test them formally in their own settings, they wish to be sure that they understand the terms and the method of scoring. The authors have found it necessary to be somewhat flexible about the coding of some items. A good example of this occurs in the German version of the HCR-20 (Müller-Isberner, Jöckel & Cabeza 1998). They call their version the 'HCR-20 + 2' to indicate that they collect somewhat more information, in a slightly different way, than is captured in the original. At a more general level, one seems to be trying as a first step to find a level of professional comfort with these devices. Always assuming that this process does not invite departure from what the scientific literature is informing one about various risk factors, there seems much to be gained from a plan of action that draws clinicians and researchers together in the common intent of making decisions that affect so vitally the lives and freedoms of persons suffering from mental and personality disorders who are in trouble with the law.

Concluding comment

This paper attempts to illustrate how the clinical practice of violence risk assessment ought to be informed by research, and how research also should be responsive to the realities of clinical practice. On the clinical side of the issue, practitioners need to be in touch with state-of-the-discipline research, demonstrate their competence in the clinical domain of risk assessment, and consider employing assessment guides. On the research side of the matter, scientists ought at least to consult with clinical colleagues, develop their research in clinically meaningful and helpful ways, and communicate this information to practitioners through clear channels.

Borum (1996) argued that '[d]espite substantive advances in knowledge about the risk for violent behavior among people with mental disorder, there have been virtually no systematic efforts to incorporate this information into a useful, empirically based framework for clinical assessment' (p. 947). He suggests that risk assessment schemes of the sort discussed in this article represent such integration and may facilitate the improvement of the clinical practice of risk assessment as well as risk assessment technology. What the present authors wish to show is that clinical practice may be flawed when it comes to risk assessment. At the same time, they wish to stress that research comes with its own limitations and liabilities. There needs to be some housecleaning on both sides of the semi-detached mansion and some thought about how best to tear down the walls which separate clinicians and researchers. This, of course, is no new issue, for it links directly to the debates around the scientist–practitioner model, and the actuarial–clinical prediction distinction, discussed earlier. It is suggested that clinical risk assessments could be better effected than at present, and that normative research could be designed so as to improve both the practice and science of violence risk assessment.

References

American Psychological Association (1992). Ethical principles of psychologists and code of conduct. *American Psychologist*, **47**, 1597–1611.

Andrews, D. A. & Bonta, J. (1995). *The psychology of criminal conduct*. Cincinnati, OH: Anderson Publishing.

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

Augimeri, L. K. & Levene, K. S. (1997). Outreach programme: Risk factors associated with possible conduct disorders and non-responders. Unpublished manuscript. Toronto, Ontario: Earlscourt Child and Family Centre.

Augimeri, L. K., Webster, C. D., Koegl, C. J. & Levene, K. S. (1998). *Early Assessment Risk List for Boys (Version 1): Consultation Edition*. Toronto, Ontario: Earlscourt Child and Family Centre.

Barratt, E. S. (1994). Impulsiveness and aggression. In J. Monahan & H. J. Steadman (Eds), *Violence and mental disorder: Developments in risk assessment*, pp. 61–79. Chicago, IL: University of Chicago Press.

Bartels, S. J., Drake, R. E., Wallach, M. A. & Freeman, D. H. (1991). Characteristic hostility in schizophrenic outpatients. *Schizophrenia Bulletin*, **17**, 163–171.

Belfrage, H. (1998). Implementing the HCR-20 scheme for risk assessment in a forensic psychiatric hospital: Integrating research and clinical practice. *Journal of Forensic Psychiatry*, **9**, 328–338.

Belfrage, H. & Fransson, G. (1997). *HCR-20: Bedömning av risk för framtida våld: Manual med instruktioner och kommentarer*. Forskningsenheten: Psykiatrist Regionvårdscentrum.

Belfrage, H., Fransson, G. & Strand, S. (1999). Prediction of violence within the correctional system using the HCR-20 risk assessment scheme. Manuscript under review.

Beutler, L. R., Williams, R. E., Wakefield, P. J. & Entwistle, S. R. (1995). Bridging scientist and practitioner perspectives in clinical psychology. *American Psychologist*, **50**, 984–994.

Binder, R. L. & McNiel, D. E. (1988). Effects of diagnosis and context on dangerousness. *American Journal of Psychiatry*, **145**, 728–732.

Binder, R. L. & McNiel, D. E. (1990). The relationship of gender to violent behavior in acutely disturbed psychiatric patients. *Journal of Clinical Psychiatry*, **51**, 110–114.

Boer, D. P., Hart, S. D., Kropp, P. R. & 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., Law, M. & Hanson, R. K. (1998). The prediction of criminal and violent recidivism among mentally disordered offenders: A meta-analysis. *Psychological Bulletin*, **123**, 123–142.

Borum, R. (1996). Improving the clinical practice of violence risk assessment. *American Psychologist*, **51**, 945–956.

Buchanan, A. (1997). The investigation of acting on delusions as a tool for risk assessment in the mentally disordered. *British Journal of Psychiatry*, **170** (s32), 12–16.

Buchanan, A., Reed, A., Wessely, S., Garety, P., Taylor, P., Grubin, D. & Dunn, G. (1993). Acting on delusions. II: The phenomenological correlates of acting on delusions. *British Journal of Psychiatry*, **163**, 77–81.

Cleckley, H. (1941). *The mask of sanity*. St Louis, MO: Mosby.

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

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*, 2nd ed. Hillsdale, NJ: Lawrence Erlbaum.

Cohen, J. (1992). A power primer. *Psychological Bulletin*, **112**, 155–159.

Convit, A., Isay, D., Gadioma, R. & Volavka, J. (1988). Underreporting of physical assaults in schizophrenic inpatients. *Journal of Nervous and Mental Disease*, **176**, 507–509.

Cooke, D. J. & Michie, C. (1997). An item response theory analysis of the Hare Psychopathy Checklist—Revised. *Psychological Assessment*, **9**, 3–14.

Cornell, D., Warren, J., Hawk, G., Stafford, E., Oram, G. & Pine, D. (1996). Psychopathy in instrumental and reactive violent offenders. *Journal of Consulting and Clinical Psychology*, **64**, 783–790.

Dempster, R. J. (1998). Prediction of sexually violent recidivism: A comparison of risk assessment instruments. Unpublished master's thesis. Simon Fraser University, Burnaby, British Columbia, Canada.

Department of Health (UK) (1990). *The care programme approach for people with a mental illness referred to the specialist psychiatry service*. London: Department of Health.

Douglas, K. S. & Hart, S. D. (1999). Psychosis as a risk factor for violence: A quantitative review of the research. Manuscript under review.

Douglas, K. S., Klassen, C., Ross, C., Hart, S. D. & Webster, C. D. (1998). *Psychometric properties of HCR-20 violence risk assessment scheme in insanity acquittees*. Poster presented at the Annual meeting of the American Psychological Association, August, San Francisco.

Douglas, K. S., Macfarlane, E. & Webster, C. D. (1996). Predicting dangerousness in the contemporary Canadian mental health and criminal justice systems. *Canada's Mental Health*, **43**, 4–11.

Douglas, K. S., Ogloff, J. R. P. & Nicholls, T. L. (1997). *The role of personality disorders in community violence among civil psychiatric patients*. Paper presented at the Fifth International Congress of the Disorders of Personality, June, Vancouver, Canada.

Douglas, K. S., Ogloff, J. R. P., Nicholls, T. L. & Grant, I. (in press). Assessing risk for violence among psychiatric patients: The HCR-20 risk assessment scheme and the Psychopathy Checklist: Screening Version. *Journal of Consulting and Clinical Psychology*.

Douglas, K. S. & Webster, C. D. (1999*a*). Predicting violence in mentally and personality disordered individuals. In R. Roesch, S. D. Hart & J. R. P. Ogloff (Eds), *Psychology and law: The state of the discipline*, pp. 175–239. New York: Plenum.

Douglas, K. S. & Webster, C. D. (1999*b*). The HCR-20 violence risk assessment scheme: Concurrent validity in a sample of incarcerated offenders. *Criminal Justice and Behavior*, **26**, 3–19.

Douglas, K. S., Webster, C. D. & Wintrup, A. (1996). *The HCR-20 Risk Assessment Scheme: Psychometric properties in two samples*. Poster presented at the annual convention of the American Psychological Association, August, Toronto, Ontario.

Dunbar, E. (1999). A psychographic analysis of violent hate crime perpetrators: Aggressive, situational and ideological characteristics of bias motivated offenders. Manuscript under review.

Estroff, S. E. & Zimmer, C. (1994). Social networks, social support, and violence among persons with severe, persistent mental illness. In J. Monahan & H. J. Steadman (Eds), *Violence and mental disorder: Developents in risk assessment*, pp. 259–295. Chicago, IL: University of Chicago Press.

Felson, R. B. (1992). 'Kick 'em when they're down': Explanations of the relationship between stress and interpersonal aggression and violence. *Sociological Quarterly*, **33**, 1–16.

Fensterheim, H. & Raw, S. D. (1996). Psychotherapy research is not psychotherapy practice. *Clinical Psychology: Science and Practice*, **3**, 168–171.

Gardner, W., Lidz, C. W., Mulvey, E. P. & Shaw, E. C. (1996*a*). A comparison of actuarial methods for identifying repetitively violent patients with mental illnesses. *Law and Human Behavior*, **20**, 35–48.

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

Goldfried, M. R. & Wolfe, B. E. (1996). Psychotherapy practice and research. *American Psychologist*, **51**, 1007–1016.

Grann, M., Belfrage, H. & Tengström, A. (in press). Actuarial assessment of risk for violence: Predictive validity of the VRAG and the historical part of the HCR-20. *Criminal Justice and Behavior*.

Grisso, T. & Appelbaum, P. S. (1992). Structuring the debate about ethical predictions of future violence. *Law and Human Behavior*, **17**, 482–485.

Grove, W. M. & Meehl, P. E. (1996). Comparative efficiency of informal (subjective, impressionistic) and formal (mechanical, algorithmic) prediction procedures: The clinical–statistical controversy. *Psychology, Public Policy, and Law*, **2**, 293–323.

Grubin, D. (1997). Predictors of risk in serious sex offenders. *British Journal of Psychiatry*, **170**, s17–s21.

Hare, R. D. (1991). *Manual for the Hare Psychopathy Checklist—Revised*. Toronto: Multi-Health Systems.

Hare, R. D. (1993). *Without conscience: The disturbing world of the psychopaths among us*. Toronto: Pocket Books.

Hare, R. D. (1996). Psychopathy: A clinical construct whose time has come. *Criminal Justice and Behavior*, **23**, 25–54.

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

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

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

Hart, S. D., Cox, D. & Hare, R. D. (1995). *The Hare Psychopathy Checklist: Screening Version (PCL:SV)*. Toronto, Ontario: Multi-Health Systems.

Hart, S. D., Webster, C. D. & Menzies, R. J. (1993). A note of portraying the accuracy of violence predictions. *Law and Human Behavior*, **17**, 695–700.

Hayes, S. C. (1996). Creating the empirical clinician. *Clinical Psychology: Science and Practice*, **3**, 179–181.

Heilbrun, K. (1997). Prediction versus management models relevant to risk assessment: The importance of legal decision-making context. *Law and Human Behavior*, **21**, 347–359.

Heilbrun, K., Philipson, J., O'Neill, M., Paninopolous, M., Strohman, L. & Bowman, Q. (1998). *Expert and practitioner approaches to communicating violence risks*. Paper presented at the annual convention of the American Psychological Association, August, San Francisco.

Hellman, D. & Blackman, J. (1966). Enuresis, firesetting and cruelty to animals: A triad predictive of adult crime. *American Journal of Psychiatry*, **122**, 1431–1435.

Hemphill, J. F., Hare, R. D. & Wong, S. (1998). Psychopathy and recidivism: A review. *Legal and Criminological Psychology*, **3**, 139–170.

Hillbrand, M. (1995). Aggression against self and aggression against others in violent psychiatric patients. *Journal of Consulting and Clinical Psychology*, **63**, 668–671.

Hodgins, S. (1983). A follow-up study of persons found incompetent to stand trial and/or not guilty by reason of insanity in Quebec. *International Journal of Law and Psychiatry*, **6**, 399–411.

Hoffman, P. B. & Beck, J. L. (1985). Recidivism among released federal prisoners: Salient factor score and five-year follow-up. *Criminal Justice and Behavior*, **12**, 501–507.

Hollander, E. & Stein, D. J. (Eds) (1995). *Impulsivity and aggression*. Toronto: Wiley.

Jackson, J. (1997). A conceptual model for the study of violence and aggression. In C. D. Webster & M. A. Jackson (Eds), *Impulsivity: Theory, assessment and treatment*, pp. 233–247. New York: Guilford.

Kanfer, F. H. (1990). The scientist–practitioner connection: A bridge in need of constant attention. *Professional Psychology: Research and Practice*, **21**, 264–270.

Klassen, D. & O'Connor, W. A. (1988*a*). Crime, inpatient admissions, and violence among male mental patients. *International Journal of Law and Psychiatry*, **11**, 305–312.

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

Klassen, D. & O'Connor, W. A. (1988*c*). A prospective study of predictors of violence in adult male mental health admission. *Law and Human Behavior*, **12**, 143–158.

Klassen, D. & 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*, **1**, 75–81.

Klassen, D. & O'Connor, W. A. (1994). Demographic and case history variables in risk assessment. In J. Monahan & H. J. Steadman (Eds), *Violence and mental disorder: Developments in risk assessment*, pp. 229–258. Chicago, IL: University of Illinois Press.

Kropp, P. R., Hart, S. D., Webster, C. D. & Eaves, D. (1999). *Manual for the Spousal Assault Risk Assessment Guide*, 3rd ed. Toronto: Multi-Health Systems.

Lidz, C. W., Mulvey, E. P. & Gardner, W. (1993). The accuracy of predictions of violence to others. *Journal of the American Medical Association*, **269**, 1007–1111.

Lindqvist, P. & Allebeck, P. (1990). Schizophrenia and crime: A longitudinal follow-up of 644 schizophrenics in Stockholm. *British Journal of Psychiatry*, **157**, 345–350.

Link, B. G. & Stueve, A. (1994). Psychotic symptoms and the violent/illegal behavior of mental patients compared to community controls. In J. Monahan & H. J. Steadman (Eds), *Violence and mental disorder: Developments in risk assessment*, pp. 137–159. Chicago, IL: University of Chicago Press.

Lipsey, M. W. & Wilson, D. B. (1993). The efficacy of psychological, educational, and behavioral treatment: Confirmation from meta-analyses. *American Psychologist*, **48**, 1181–1209.

Lyon, D. R., Hart, S. D. & Webster, C. D. (in press). Violence and risk assessment. In R. Schuller & J. R. P. Ogloff (Eds), *Law and psychology: Canadian perspectives*. Toronto: University of Toronto Press.

MacCulloch, M. J., Snowden, P. R., Wood, P. J. W. & Mills, H. E. (1983). Sadistic fantasy, sadistic behaviour, and offending. *British Journal of Psychiatry*, **143**, 20–29.

McNiel, D. E. & Binder, R. L. (1989). Relationship between preadmission threats and later violent behavior by acute psychiatric inpatients. *Hospital and Community Psychiatry*, **40**, 605–608.

McNiel, D. E. & Binder, R. L. (1991). Clinical assessment of the risk of violence among psychiatric inpatients. *American Journal of Psychiatry*, **148**, 1317–1321.

McNiel, D. E. & Binder, R. L. (1994*a*). The relationship between acute psychiatric symptoms, diagnosis, and short-term risk of violence. *Hospital and Community Psychiatry*, **45**, 133–137.

McNiel, D. E. & Binder, R. L. (1994*b*). Screening for risk of inpatient violence: Validation of an actuarial tool. *Law and Human Behavior*, **18**, 579–586.

McNiel, D. E. & Binder, R. L. (1995). Correlates of accuracy in the assessment of psychiatric inpatients' risk of violence. *American Journal of Psychiatry*, **152**, 901–906.

McNiel, D. E., Binder, R. L. & Greenfield, T. K. (1988). Predictors of violence in civilly committed acute psychiatric patients. *American Journal of Psychiatry*, **145**, 965–970.

Meehl, P. E. (1954). *Clinical versus statistical prediction*. Minneapolis, MN: University of Minnesota Press.

Meehl, P. (1997). Credentialed persons, credentialed knowledge. *Clinical Psychology: Science and Practice*, **4**, 91–98.

Menzies, R. & Webster, C. D. (1995). Construction and validation of risk assessments in a six-year follow-up of forensic patients: A tridimensional analysis. *Journal of Consulting and Clinical Psychology*, **63**, 766–778.

Menzies, R. J., Webster, C. D. & Sepejak, D. S. (1985). Hitting the forensic sound barrier: Predictions of dangerousness in a pre-trial psychiatric clinic. In C. D. Webster, M. H. Ben-Aron & S. J. Hucker (Eds), *Dangerousness: Probability and prediction, psychiatry and public policy*, pp. 115–143. New York: Cambridge University Press.

Monahan, J. (1981). *Predicting violent behavior: An assessment of clinical techniques*. Beverly Hills, CA: Sage.

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

Monahan, J. (1988). Risk assessment of violence among the mentally disordered: Generating useful knowledge. *International Journal of Law and Psychiatry*, **11**, 249–257.

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

Monahan, J. (1996). Violence prediction: The last 20 and the next 20 years. *Criminal Justice and Behavior*, **23**, 107–120.

Monahan, J. & Steadman, H. J. (Eds) (1994*a*). *Violence and mental disorder: Developments in risk assessment*. Chicago, IL: University of Chicago Press.

Monahan, J. & Steadman, H. J. (1994*b*). Toward a rejuvenation of risk assessment research. In J. Monahan & H. J. Steadman (Eds), *Violence and mental disorder: Developments in risk assessment*, pp. 1–17. Chicago, IL: University of Chicago Press.

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

Monahan, J., Steadman, H. J., Appelbaum, P. S., Robbins, P. C., Mulvey, E. P., Silver, E., Roth, L. H. & Grisso, T. (in press). Developing a clinically useful actuarial tool for assessing violence risk. *British Journal of Psychiatry*.

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

Müller-Isberner, J. R. & Jöckel, D. (1997). *The implementation of the HCR-20 in a German hospital order institution*. Paper presented at the Seventh European Conference on Psychology and Law, September, Solna, Sweden.

Müller-Isberner, R., Jöckel, D. & Cabeza, S. G. (1998). *Die Vorhersage von Gewalttaten mit dem HCR 20*. Haina: Institut für Forensische Psychiatrie Haina.

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

Mulvey, E. P., Shaw, E. & Lidz, C. W. (1994). Why use multiple sources in research on patient violence in the community? *Criminal Behaviour and Mental Health*, **4**, 253–258.

Nezu, A. M. (1996). What are we doing to our patients and should we care if anyone else knows? *Clinical Psychology: Science and Practice*, **3**, 160–163.

Nicholls, T. L., Ogloff, J. R. P. & Douglas, K. S. (1997). Comparing risk assessments with female and male psychiatric outpatients: Utility of the HCR-20 and Psychopathy Checklist: Screening Version. In J. R. P. Ogloff (Chair), *Involuntary civil commitment—patient characteristics, review panel decision making, and risk assessment*. Symposium conducted at the annual convention of the American Psychological Association, August, Chicago.

Novaco, R. W. (1994). Anger as a risk factor for violence among the mentally disordered. In J. Monahan & H. J. Steadman (Eds), *Violence and mental disorder: Developments in risk assessment*, pp. 21–59. Chicago, IL: University of Chicago Press.

Nuffield, J. (1982). *Parole decision-making in Canada: Research towards decision guidelines*. Ottawa: Ministry of Supply and Services Canada.

Overall, J. E. & Klett, C. J. (1962). The brief psychiatric rating scale. *Psychological Reports*, **10**, 799–812.

Pfohl, S. J. (1978). *Predicting dangerousness: The social construction of psychiatric reality*. Lexington, MA: Lexington Books.

Quinsey, V. L., Harris, G. T., Rice, G. T. & Cormier, C. A. (1998). *Violent offenders: Appraising and managing risk*. Washington, DC: American Psychological Association.

Reed, J. (1997). Risk assessment and clinical risk management: The lessons from recent inquiries. *British Journal of Psychiatry*, **170**, s4–s7.

Rice, M. E. (1997). Violent offender research and implications for the criminal justice system. *American Psychologist*, **52**, 414–423.

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

Ross, D. J., Hart, S. D. & Webster, C. D. (1998). *Aggression in psychiatric patients: Using the HCR-20 to assess risk for violence in hospital and in the community*. Port Coquitlam, British Columbia: Riverview Hospital.

Salekin, R. T., Rogers, R. & 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*, **3**, 203–215.

Saunders, D. G. (1992). Woman battering. In R. T. Ammerman & M. Hersen (Eds), *Assessment of family violence: A clinical and legal sourcebook*, pp. 208–235. New York: Wiley.

Schopp, R. F. (1996). Communicating risk assessments: Accuracy, efficacy, and responsibility. *American Psychologist*, **9**, 939–944.

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

Singer, J. L. (1990). The scientific basis of psychotherapy practice: A question of values and ethics. *Psychotherapy: Theory, Research, and Practice*, **17**, 372–383.

Snowden, P. (1997). Practical aspects of clinical risk assessment and management. *British Journal of Psychiatry*, **170**, s32–s34.

Sobell, L. C. (1996). Bridging the gap between scientists and practitioners: The challenge before us. *Behavior Therapy*, **27**, 297–320.

Steadman, H. J., Monahan, J., Appelbaum, P. S., Grisso, T., Mulvey, E. P., Roth, L. H., Robbins, P. C. & Klassen, D. (1994). Designing a new generation of risk assessment research. In J. Monahan & H. J. Steadman (Eds), *Violence and mental disorder: Developments in risk assessment*, pp. 297–318. Chicago, IL: University of Chicago Press.

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

Steadman, H. J., Silver, E., Monahan, J., Appelbaum, P. S., Robbins, P. C., Mulvey, E. P., Grisso, T., Roth, L. H. & Banks, S. (in press). *Law and Human Behavior*.

Strand, S., Belfrage, H., Fransson, G. & Levander, S. (1999). Clinical and risk management factors in risk prediction of mentally disordered offenders: More important than actuarial data? *Legal and Criminological Psychology*, **4**, 67–76.

Stricker, G. (1992). The relationship of research to clinical practice. *American Psychologist*, **47**, 543–549.

Stricker, G. & Trierweiler, S. J. (1995). The local clinical scientist: A bridge between science and practice. *American Psychologist*, **50**, 995–1002.

Swanson, J. W. (1994). Mental disorder, substance abuse, and community violence: An epidemiological approach. In J. Monahan & H. J. Steadman (Eds), *Violence and mental disorder: Developments in risk assessment*, pp. 101–136. Chicago, IL: University of Chicago Press.

Swanson, J. W., Borum, R., Swartz, M. & Monahan, J. (1996). Psychotic symptoms and disorders and the risk of violent behavior in the community. *Criminal Behaviour and Mental Health*, **6**, 317–338.

Webster, C. D. & Cox, D. N. (1997). Integration of nomothetic and ideographic positions in risk assessment: Implications for practice and the education of psychologists and other mental health professionals. *American Psychologist*, **52**, 1245–1246.

Webster, C. D., Dickens, B. M. & Addario, S. M. (1985). *Constructing dangerousness: Scientific legal and policy implications*. Toronto, Ontario: University of Toronto Centre of Criminology.

Webster, C. D., Douglas, K. S., Eaves, S. D. & Hart, S. D. (1997*a*). Assessing risk of violence to others. In C. D. Webster & M. A. Jackson (Eds), *Impulsivity: Theory, assessment, and treatment*, pp. 251–277. New York: Guilford.

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

Webster, C. D., Eaves, D., Douglas, K. S. & Wintrup, A. (1995). *The HCR-20 scheme: The assessment of dangerousness and risk*. Vancouver, B.C.: Mental Health Law and Policy Institute, and Forensic Psychiatric Services Commission of British Columbia.

Webster, C. D. & Jackson, M. A. (Eds) (1997). *Impulsivity: Theory, assessment, and treatment*. New York: Guildford.

Wessely, S. C., Castle, D., Douglas, A. J. & Taylor, P. T. (1994). The criminal careers of incident cases of schizophrenia. *Psychological Medicine*, **24**, 483–502.

Wintrup, A. (1996). Assessing risk of violence in mentally disordered offenders with the HCR-20. Unpublished master's thesis. Simon Fraser University, Burnaby, British Columbia, Canada.