Northumbria Research Link

Clinical judgement in violence risk assessment

Jennifer Murray
Glasgow Caledonian University

Dr. Mary E. Thomson
Glasgow Caledonian University

Abstract

The present article discusses the three main approaches to violence risk assessment, clinical judgement, actuarial assessment, and structured clinical judgement, informing the reader of the comparative benefits and short-comings of these methods of violence risk assessment. In particular, the present article highlights the controversy within the literature surrounding clinical judgement in comparison to actuarial assessments of violence risk, and proposes that the statistically significant 'improvements' of violence prediction when using actuarial scales in comparison to clinical predictions of dangerousness do not necessarily measure the skill of the clinician adequately. Specifically, an assessment of 'dangerousness' does not equal a prediction of violent recidivism. It is argued that clinicians are not predictive forecasters of risk, as in actuarial scales, but are, rather, trained to manage risk. In addition, suggestions for future research directions in the field of improving violence risk assessments are made.

Keywords – clinical judgement, actuarial assessment, structured clinical judgement, violence, violence risk assessment.

Within the field of violence risk assessment, clinicians are reported to be increasing their use of, and are indeed being encouraged to use, standardised risk assessment tools to supplement their own judgements (Scott & Resnick, 2006). Historically, clinicians would be expected to conduct violence risk assessments based upon their own experience in practice using judgement alone, without the use of standardised aids (Harris & Lurigio, 2007). While this approach is still common in practice (Cooke, Michie, & Ryan, 2001), increasing concerns over the accuracy of clinical judgements of violence risk assessment (for example, Monahan, 2001) have led to a vast number

of tools and approaches being developed regarding the way in which violence risk assessments should be conducted. Three key approaches to the way in which violence risk assessments are conducted exist: clinical judgement alone; actuarial scales, which are standardised measures based on statistical models used to predict the risk of future violence; and structured assessment tools to aid clinical judgement, which allow clinicians to use their experience to make judgements in a clear and structured manner using standardised guidelines. However, there is still a great deal of debate within the academic literature regarding the most effective approach to conducting violence risk assessments. For example, while Quinsey, Harris, Rice, and Cormier (1999) suggested that clinical judgement should be replaced completely by the use of actuarial scales, Douglas, Ogloff, and Hart (2003) have criticised actuarial tools for being rigid, lacking sensitivity to change and failing to aid risk management and the prevention of violence.

In relation to clinical judgement, the debate centres on the accuracy of the clinician's ability to predict violence. Although clinical judgement is reported as being the most commonly used method in clinical practice (Cooke *et al.*, 2001), this approach has been widely criticised as inaccurate. In fact, Monahan's (1984) claim that two thirds of all clinical predictions of violence are incorrect has been widely cited. In view of such claims, the present paper will discuss each of the three key methods of risk assessment, highlighting their strengths and weaknesses, thus providing the reader with an adequate knowledge of the current situation surrounding the practice of violence risk assessment. In order to achieve this, however, the wider ranging importance of clinical accuracy in violence risk assessments must first be understood.

The accuracy of clinical judgements has been a highly researched area since Ashley's (1922) research, which investigated whether released psychiatric patients posed a real criminal threat to the public. In this research, the criminal behaviour of 700 psychiatric patients was followed for three months after their release from hospital. Ashley (1922) reported that only twelve of these patients had been re-arrested for a variety of violent and non-violent crimes. While this finding would seem to indicate a poor relationship between mentally ill individuals and criminal behaviour, Harris and Lurigio (2007) highlighted that the findings of Ashley's research cannot be interpreted with high levels of confidence as the arrest rates for the patients and the general population were never compared. Similarly, Thornberry and Jacoby (1979) followed up the rates of violent offences committed by individuals who had been released from hospitals for the criminally insane to civil hospitals as a result of judicial decisions. While these individuals had previously been

deemed as too dangerous for release, of the 65% who were subsequently released into the community, only 11% were re-arrested for violent offences.

Studies such as Ashley's (1922) and Thornberry and Jacoby's (1979) would appear to indicate that mentally ill individuals in the past were unjustly confined due to faulty judgements assessing an above average level of risk of dangerousness which resulted in them being confined within facilities which were more restrictive than was perhaps necessary (Litwack, Zapf, Groscup, & Hart, 2006). The long standing popularity of this field of research can therefore be attributed to its multi-faceted appeal and applications: it is important to assess the risk of violence accurately from both a legal and an ethical standpoint. In relation to the former application, clinicians are called upon explicitly to conduct violence risk assessments within a legal setting (Dempster, 2003). The apparent increase in the legal requirement of risk assessments has been attributed to a number of factors, including changes in the law relating to mentally ill offenders (for example, the Baxtrom Ruling, 1966, which led to 967 offenders being released or transferred from high security to low security psychiatric hospitals [Steadman & Cocozza, 1974]) in addition to the increase in media portrayal of violent crimes being attributed to mentally ill individuals and the subsequent public concerns and fear brought about by these factors (Douglas & Webster, 1999). It has therefore been predicted that concerns relating to mentally ill individuals committing violent acts will remain a core issue in both mental health law and academic research alike (Monahan, 1988).

The key focus of recent research relating to the field of violence risk assessment centres on the best methods used in predicting future violence, i.e., through using a clinical or actuarial approach. This comparative vein of research has been popular largely due to the demands placed on clinicians to produce accurate and reliable risk assessments of future violence. There are now a large number of violence risk assessment tools available to the clinician when conducting a risk assessment. The particular tool chosen will depend on a number of factors, including the clinician's preferred tool, the most suitable tool for the specific case (e.g., the VRAG for predicting recidivism in male violent offenders and its derivative scale, the SORAG, for assessing male sexual offenders [Quinsey, Harris, Rice, & Cormier, 1998], the HCR-20 for assessing risk of violence in a structured manner across a broad range of settings [Webster, Douglas, Eaves, & Hart, 1997] and, more recently, the development of the PRISM for assessing situational/institutional violence [Johnstone & Cooke, 2008]) and which tool the funding body or employer uses or supports.

Historically, clinicians would be called upon to use their own experience and intuition when conducting violence risk assessments. This method, while still common in

practice, has been highly criticised as being limited in terms of both accuracy and inter-clinician agreement (Hart, Michie, & Cooke, 2007). While it has now become common within the literature to find statements which suggest that actuarial assessments are superior to unstructured clinical assessments, there is evidence to suggest that this may not be the case (Litwack, 2001). For example, in an analysis of the statistical and clinical evidence, Greenland (1980) found that psychiatrists may have been more accurate in their prediction of violence than previously reported when the type of offender is considered and controlled. Furthermore, through the utilisation of clinical experience when making judgements about the risk of violence posed by an individual, the clinician can assess emotional state throughout the interview. For example, emotional traits such as lack of empathy and anger (Menzies, Webster, & Sepejak, 1985) and physiological and behavioural traits such as chanting, flared nostrils, flushed face and clenching of the jaw and hands (Berg, Bell, & Tupin, 2000) are among the indicators of violence that may not be detected through the use of actuarial tools alone. The use of both actuarial tools and of structured risk assessment instruments in conjunction with clinical interviews has therefore been judged to be most effective (Scott & Resnick, 2006). This recommendation harvests the benefits of both actuarial tools and clinical judgement (structured or unaided) while minimizing their limitations in practice.

The VRAG is the most widely used actuarial scale for predicting violence (Cooke *et al.*, 2001). This tool was developed by Quinsey *et al.* (1998) in order to predict the likelihood of violent behaviour enacted by institutionalised adult males. Although the predictive accuracy of the VRAG has been widely demonstrated, it has also come under criticism. Critique aimed specifically at the VRAG highlights the issue of insensitivity to the type of violence committed: the level of severity of the violent act being assessed is not taken into consideration or weighted differently by the VRAG, and, as such, important information about the offence may be lost. For instance, when using the VRAG the violent act of murder would carry the same weight as that of assault, despite the much greater severity of the consequences to the victim. In addition, the VRAG does not take into account drug or substance abuse as a risk factor for violence, despite this factor being a key factor in predicting recidivism (Howard, 2009). No specific reason is given for the omission of this important risk factor from the VRAG; however, it is logical that the omission would be based on the characteristics of the base sample. That is, based on the sample on which the VRAG was developed, drug or substance abuse may not have been a common factor associated with violence, despite this being a key predictive factor in other offending populations. Thus, while the information gathered through this tool may be of use in predicting risk, there is a possibility of missing important 'broken leg' or dynamic risk cues. These so called 'broken leg' cues are changeable in nature and

over time, with the classic example of an individual having a broken leg: in the short term, this would likely hinder the individual's opportunity to engage in violent behaviour, thus lowering their level of risk and dramatically altering their risk management needs. Thus, in this instance the clinician may choose to adjust the actuarial prediction to take into account this dynamic risk/protective factor. A similar adjustment may also be made in the case of drug abuse (which, as mentioned is not included in the VRAG), thus aiding the formulation of the risk management plan while still taking into account the important risk factors highlighted within the actuarial tool.

This cautiously adjusted actuarial approach is not, however, ideal (Kriegman, 2007), as clinicians tend to identify too many exceptions to the actuarial formula, thus modifying the prediction to a great extent (Dawes, Faust, & Meehl, 1989); however, Dawes *et al.* (1989) suggested that the inclusion of the clinician's modifications does not necessarily aid the prediction, and if over-adjusted may, in fact, negatively affect the accuracy of the prediction. In addition, practical limitations of time may not always allow for such 'full' and rounded assessments.

While actuarial tools have been widely criticised for being rigid, lacking sensitivity to change and for not being generalisable to populations out with the sample on which they were based (Scott & Resnick, 2006), it is thought that these tools can be of great use to the clinician when gathering information in the initial assessment stages, since many of the factors included within actuarial assessments are those upon which clinical assessments are based (Litwack *et al.*, 2006).

In common with clinical judgement, however, actuarial tools are limited when used alone. Within true actuarial assessments, no clinical judgement is present but instead strict, uniform rules and decision making criteria are adhered to. It is largely considered that due to the statistical improvement in inter-rater reliability (Dempster, 2003), that actuarial assessments of violence risk are superior to clinical judgements. In fact, Quinsey *et al.* (1999) have recommended the complete replacement of clinical evaluations of dangerousness with actuarial tools. However, concerns relating to the highly variable nature of forensic cases have led to criticisms of this recommendation: because cases are diverse in nature and because it would be impossible to mathematically model every relevant clinical factor, to rely solely on an actuarial tool could cause potentially important indicators of violent behaviour to be missed.

The use of structured clinical judgement has been seen as a step towards more accurate decision making practice, with this approach being based upon

combining empirically established risk factors with clinical judgement (Douglas *et al.*, 1999). This method of violence risk assessment is considered advantageous as the clinician is given some level of flexibility and discretion. In addition, working to non-weighted guidelines that are easily adapted to new cases (Hart, 2000) allows the decision making process to be recorded. The factors on which the final judgement is made are explicit, as in actuarial tools, thus the level of inter-rater reliability can be considered as higher than would be the case with clinical judgement alone. The HCR-20 is one of the most widely researched structured risk assessment guides and has been boasted as "the leading instrument of the structured professional judgement approach" (Lamont & Brunero, 2009). This broadband risk assessment measure was developed by Webster *et al.* (1997) to assess violence across a range of situations, and has been validated in both forensic and civil psychiatric populations (Canter & Zukauskiene, 2008). It comprises 20 items across three subscales: 10 Historical items, 5 Clinical items and 5 Risk Management items. Unlike the actuarial tools of the past, the HCR-20 (and structured risk assessment tools in general) do not seek to predict violence, but instead aim to identify the presence of relevant risk factors, to assess and indicate the individual's current level of risk posed, and to inform suitable interventions in order to best manage the presenting level of risk. Thus, rather than predicting the risk of future violence that an individual poses, the HCR-20 instead acts to inform and manage this risk.

In a similar manner as with actuarial tools, it has been reported that these structured tools are also being used in an 'adjusted' manner, with the guidelines proposed by these tools not being applied uniformly; instead, these guidelines are reported as being used in a selective manner, with expert users substituting their own techniques in individual cases (Garfield & Garfield, 2000). This selective use, while not entirely in keeping with 'evidence-based practice' may perhaps represent one of the key attributes of violence risk assessment: each case is individual. As instances of violent behaviour can be highly individual in nature, one must take into account that no single evidence-based risk assessment tool or set of guidelines can capture every aspect of every case. As such, the experienced clinician may choose to supplement the risk assessment guideline with additional knowledge about the case that they feel is of paramount importance in assessing the level of risk posed. While this was not an easy option when using the earlier actuarial risk assessment tools which were weighted using mathematical formula (e.g., the VRAG), structured risk assessment guidelines (such as the HCR-20) do allow the clinician some freedom to include additional information that they feel is relevant to the case, as the completion of the assessment is not only informed by historical or case file information, but also through clinical interviews thus preventing important 'broken leg' information from being 'lost' through the process of assessing risk.

Despite the controversy relating to the accuracy of clinician judgement (Borum, 1996) and the reliability of actuarial tools used within violence risk assessments (Litwack, 2001), clinicians continue to be called upon within a courtroom setting to assess whether or not an individual poses a serious level of risk to the public (Litwack *et al.*, 2006), as previously discussed. Typically, clinicians in these circumstances are asked to assess the likelihood of an individual committing violent behaviour if released into the public (Rabinowitz & Garelik-Wyler, 1999). However, within mental and forensic health services, clinicians are not primarily trained in behavioural forecasting but are instead trained in providing care for their patients and taking preventative action where vulnerable individuals may pose a risk of harming themselves and/or others (Rabinowitz & Garelik-Wyler, 1999). It must be noted that a clinician's assessment of an individual as 'dangerous' is not a guarantee or a prediction of violent behaviour, but is, rather, an indication that the individual poses a significant risk of committing a violent act in the future. As illustrated by Litwack *et al.* (2006), while a patient may be assessed by a clinician as too dangerous for release, this individual may not commit violent acts when released. This cannot be seen as a faulty judgement on the part of the clinician, who merely acted within their boundaries of responsibility in assessing the individual's overall level of risk and was therefore not necessarily acting within a behavioural forecasting role. In addition, while much of the research relating to clinical accuracy in the prediction of violent behaviour discusses judgements of 'dangerous' and 'non-dangerous', differences in terminology must be considered. For example, in Kozol, Boucher, and Garofalo's (1972) research investigating the accuracy of clinical judgements, individuals who were reported to be 'dangerous' and who were therefore considered within the research to be likely to commit violent acts were, instead, simply *not recommended for release* by clinicians.

There is clearly a distinction between an individual being not recommended for release and being directly estimated to commit a violent act. The terminology on which dangerous and non-dangerous groups are categorised prior to making bold statements pertaining to the accuracy of clinician judgements must therefore be established within specific research. For example, Lidz, Mulvey, and Gardner (1996) found that mental health professionals performed at a level that was better than chance when asked to rate cases on a scale directly assessing the clinician's opinion of the 'potential for violence' over a six month period. The authors further asserted, however, that due to considerable numbers of individuals who had been rated as having a low potential for violence committing violent acts (a measure of sensitivity) and a considerable proportion of those deemed as having a high potential for violence not committing violent acts (a measure of specificity) within the six month follow up period, that clinical judgements did not have a high enough

level of sensitivity or specificity to produce accurate predictions of violence. The terminology used in this study must again be considered, as rating the 'potential for violence' is once again not a pure prediction of violence. As illustrated, it is clear that the findings of research relating to violence risk assessment are highly inconsistent, with Monahan (1988) going as far as to state that for every piece of research that may find an increase in the predictive accuracy of clinicians, another will find that clinicians perform at a level that is no more accurate than chance.

Thus, in relation to the body of research concerning the accuracy of clinical judgements in violence risk assessment, it is of the utmost importance to take into account the differing contexts and definitions of 'risk' across different disciplines. The concept of risk encompasses "the nature, severity, imminence, and frequency or duration of harm – as well as its likelihood" (Litwack *et al.*, 2006, pp.493). Risk can therefore be considered as dynamic, with Hart (2000) further adding that risk, when considered in terms of violence prediction, is an unknown hazard that can must be predicted with a level of uncertainty. From these definitions, it can clearly be extruded that there is no single-faceted definition of risk in terms of clinicians' assessments of violence risk and prediction.

It is for this reason that much of the previous research on clinician accuracy has since been criticised. For example Kozol *et al.* (1972) conducted research into the accuracy of clinical judgements relating to predicting future violence. They followed up a cohort of males who had been released from restrictive facilities over a five year period. The sample consisted of 592 males who had been convicted for assaultive offences and who had been assessed by a team of mental health professionals using clinical interviews, life histories and psychological tests. 386 of the sample were released on the basis of being assessed as no longer dangerous, with a further 49 being released against the advice of a mental health team. At the conclusion of the five year follow up period, Kozol *et al.* (1972) reported that only 8% of the non-dangerous sample recidivised, while 34.7% of those who had been diagnosed as dangerous were found to have committed a serious crime.

The findings of Kozol *et al.'s* (1972) research have been widely cited as a demonstration of clinical predictions of violence being incorrect every two out of three instances (e.g., Monahan, 1984). However, while it is reported that clinicians provided correct judgements of violence prediction in only 34.7% of the dangerous group, Kozol, Boucher, and Garofalo (1973) later reported that of the 49 dangerous patients released, at least fourteen would not have been diagnosed as dangerous had they been assessed in the latter half of the research due to the improved diagnostic techniques used. In addition, it must always be considered with research

reporting and measuring crime rates that criminal and recidivistic behaviour may indeed be higher than recorded. Indeed, research by Hall (1982) indicated that only 20% of serious crime may result in arrest. This is of utmost importance in assessing the level of accuracy in clinical predictions of violent behaviour when considering research of the nature of that by Kozol *et al.* (1972). It must therefore be taken into account that apparently faulty judgements may, at least in part, be the result of inaccurate measures of recording criminal behaviour. This can be evidenced by the findings of Klassen and O'Conner (1987): the authors found that by simply including a measure of self-reported violence to that of arrest rates, the accuracy of clinical predictions of violence could be increased by up to 27.8%.

Kozol *et al.'s* (1972) research does, however, suggest that clinicians are significantly better at judging accurately cases of non-violent future behaviour than predicting future violence. In addition, research by Rabinowitz and Garelik-Wyler (1999) found that psychiatric trainees staffing a psychiatric emergency room were better at predicting non-violence than predicting violence in newly admitted patients. This finding is a positive indicator in light of the concerns raised relating to the possibility of excessively high levels of false-positive judgements resulting in individuals being held in highly restrictive facilities (Litwack *et al.*, 2006). That is, in light of the afore-mentioned research findings, that the majority of the individuals who had been deemed too dangerous for release had not actually violently re-offended upon release, the findings of Kozol *et al.* (1972) and Rabinowitz and Garelik-Wyler (1999) go some way to alleviate the ethical concerns raised by the earlier research.

There are clear ethical issues associated with the high levels of positive predictions of violence: if the majority of individuals who have been judged to be 'dangerous' and have therefore not been recommended for release have been shown statistically to not behave in a dangerous manner when released, should these individuals be incarcerated in the first place? In order to answer this question, a number of factors must be considered. These shall briefly be discussed, but it must be noted that the full range of ethical issues associated with violence risk assessment and detention cannot be covered in the scope of this article. In terms of both a legal and an ethical standpoint, the Criminal Justice Mental Health Standards (Standards 7-3.10; American Bar Association, 1989) outline that an individual must have not only sufficient education and clinical training and experience, but must also have adequate forensic knowledge gained through specialized training or an alternative route in order for an individual to be court appointed to perform a forensic evaluation. The individual must, therefore, be adequately knowledgeable in both the required clinical matters and the associated legal matters prior to passing a judgement on the level of violence that an individual poses. In addition, when

assessing the level of risk that an individual poses, the expert must consider not only the individual being assessed (i.e., the offender) and the level of risk to the public, but also under what circumstances this level of risk will be best managed and in turn reduced. Thus, where an expert may recommend that an individual not be released from detention, the decision is based on not only concern for safeguarding the public but also will consider what is best for the offender in terms of lowering their likelihood of re-offending through the implementation of interventions. In addition, when considering issues related to re-offending, one must consider that the aim of an intervention, whether it is through the implementation of a clinical risk management plan or simply a period of incarceration, is to reduce the level of risk posed by an individual (Heilbrun, 1997). In contrast, the aim of the majority of published violence risk assessment research concerns the accuracy of clinical predictions (Carroll, 2007). If one considers that through the course of incarceration an individual may receive some form of therapeutic intervention (e.g., an anger management course), it then becomes somewhat intuitive that the level of risk posed will be lower at the end of the intervention/incarceration than that assigned beforehand. Thus, if the intervention has been successful a negative relationship between the original prediction made (e.g., highly dangerous) and the likelihood of re-offending *should* be found.

Furthermore, if clinicians are indeed efficient at predicting non-violence accurately, as would be suggested by Kozol *et al.*'s (1972) and Rabinowitz and Garelik-Wyler's (1999) findings, there may be fewer 'non-dangerous' individuals, where the definition of dangerousness acts as a prediction of violence, incarcerated than would be suggested when taking solely into account the earlier research findings reporting high rates of false positive predictions of violence in those incarcerated individuals (e.g., Ashley, 1922; Steadman & Cocozza, 1974; Thornberry and Jacoby, 1979). As can be seen, the issues surrounding the prediction of violent re-offending behavior found in the literature and the reality of offender reform and detention are complex. In order to judge effectively whether the high level of false positive predictions of future violence found in the literature is indeed that, or whether it is actually based on the success of the interventions applied to an individual, research which takes into account not only predictions of future violence and follow-up re-offending data, but also whether or not the individual had received any interventions must be conducted.

Rabinowitz and Garelik-Wyler (1999) further emphasise this within their violence prediction research, stating that some form of preventative action or treatment may have affected the outcome of the target individual's behaviour. That is, where a prediction of violent behaviour has been made, any treatment or therapy received

either while incarcerated or in the community may reduce the potential for violence. Thus, there has been a move within the literature which is driving the focus towards the management of risk, rather than the prediction of future violence. The very nature of violence risk assessment has come to be considered as an ongoing process, not a single standing assessment (Davidson, 1997). This movement towards viewing risk assessment as an ongoing process relating to the management of the posing risk, rather than making a one-off assessment bases on static factors, is in line with the argument initiated by Davidson (1997); that the risk of violence posed by an individual is a dynamic factor and as such is subject to change over time. Thus, rather than assessing the likelihood of risk, it is now accepted that risk must be managed.

While great attention within the literature has been given to the various methods of improving the accuracy of clinician judgements, less detail has been paid to the actual processes used by clinicians whilst assessing violence risk in practice (Elbogen, 2002) and the ways in which risk is managed in practice (Dempster, 2003). In relation to this descriptive research, Elbogen (2002) reviewed three key areas of clinician judgement: cue utilisation (the risk factors utilised in decision making), clinical reliability and clinical decision making.

In relation to cue utilisation, Elbogen's (2002) review of the descriptive research relating to clinician judgements of violence risk assessment found that clinicians tend to over-emphasise the importance of certain crimes and under-emphasise others (e.g., murder and sexual offence, respectively). It should be noted however, that it is not the primary aim of the clinician to predict future violence but rather it is to manage the risk of violent re-offending (Litwack *et al.*, 2006), as previously discussed.

Elbogen and Huss (2000) further specified that the context in which violence risk assessments are made may affect the types of risk factors focused upon when decision making. For example, Menzies and Webster (1995) found that in a forensic setting, clinicians focused their decision making on previous violence, alcohol abuse and low anger control to a greater extent than other presenting cues. Conversely, Quinsey (1995) suggested that clinicians rely on salient cues, such as extreme delusions, when assessing the risk of violence, even when these cues are not necessarily related to the case or indeed relevant. Homan, van Knippenberg, Van Kleef, and De Dreu (2007) conducted research concerning the effects of differing levels of salience in group decision making which may be relevant when considering clinical decision making and violence risk assessments conducted within group settings. In addition to levels of informational diversity to which the group was exposed, Homan *et al.* (2007) further emphasised the importance of the salience of

social categories within the group. The authors found that with increasing diversity in the social categories included within the group, increased conflict arose as a result of the formation of social subgroups. This finding is theoretically considered to arise due to the difficulties experienced by groups when faced with exploring and accepting the perspectives of several subgroups (Bowers, Pharmer, & Salas, 2000). Conversely however, van Knippenberg, De Dreu, and Homan (2004) found that diversity of information within a group can lead to elaboration of task-relevant information, arguing that this increased exchange of information will aid the accuracy of the decision making process.

These apparently opposing standpoints may, however, be considered as complimentary: in order for informational diversity to be exchanged and discussed by a group, sub-group categorisation must be minimal (Homan *et al.*, 2007). Thus, in order to attain maximal benefits within group decision making, informational diversity must be high and group heterogeneity must be minimal. However, while the impact of group processes on decision making in violence risk assessment would appear to be extremely relevant within a clinical setting due to the increasing tendency for clinicians to conduct violence risk assessments in groups (Croskerry, 2005), research of this nature (that is, the exploration of social salience/subgroup categorization on group decision making) has largely not been empirically applied to the field of clinical decision making in violence risk assessment. Thus, little is actually known about the possible effects that group decision making may have on judgements made concerning violence risk assessments. It should therefore be noted that the research discussed here regarding group decision making in violence risk assessment does not and cannot cover every theory or model of group decision making, and indeed there is an entire literature on the topic of group decision making in contexts other than violence risk assessment that may be applicable to the field. Future research investigating group processes in violence risk assessment, particularly investigating the impact of multi-disciplinary teams on judgement and decision making in this context, would therefore be beneficial.

While it is recognised that not all of the aspects of group decision making that may affect judgement in violence risk assessment can be covered in the scope of the current article, a number of the main factors shall now be discussed. For example, false congruence in decision making may occur as a result of group processes. It has been well documented that in group situations the diversity of information and perspectives is often not exploited to full potential (Bowers *et al.*, 2000; Homan *et al.*, 2007; Webber & Donahue, 2001). While it is thought that group decision making may be better than individual decision making due to the ability to pool unique information (Hollingshead, 1996), it has been found that group discussions

disproportionately reflect information that is know by all members of the group; thus making information that is known by only one member of the group less discussed and therefore less important in terms of weighting in the final decisions made by the group (Hollingshead, 1996). This may then lead to a false sense of congruence in the final decision made and result ultimately lead to a less than optimal decision being made (Stasser & Titus, 1985). However, Stasser, Stewart, and Wittenbaum (1995) found that the heightened influence of shared knowledge compared to unique knowledge in group settings could be decreased by making it clear to all members of the group that each member has additional, non-shared information. With regard to violence risk assessment, when discussing a case the multi-disciplinary team should explicitly recognise the value of individual information that may not be held by all of the group members or indeed contained within the violence risk report or case being discussed.

In addition, Hellingshead (1996) highlighted that through having access to information profiles rather than having to rely on memory alone, groups discussed both unique and common information to a greater extent. However, when asked to write rationales for their decisions, groups who had access to the information profiles did not have a higher mention of the unique information when compared to groups who relied on memory alone. Stewart and Stasser (1995) also found that unique information was omitted from recalled information where it was previously included in discussion, with the exception of when the unique information was presented by an individual who was considered to be an expert. Homan *et al.* (2007) argued that it is not only the diversity of information and the way that this is shared within a group that may affect decision making, but also the diversity of the group members. This has major implications in terms of the decisions reached by multi-disciplinary teams in violence risk assessment. A recent study by Kettles, Gass, Addo, and Mckie (2009) found that key care decisions within intensive psychiatric care units still tended to be made by doctors and nurses, and that decisions made by medical staff held more authority than those made by other members of the care teams. Kettles *et al.* (2009) concluded from these findings that inequalities in power still exist within multi-disciplinary teams. With a related focus on leadership in group decision making, Taggar and Ellis (2007) conducted research which investigated the role of leaders in shaping team norms, based on Feldman's (1984) suggestion that team leaders have a strong impact on team norm development. It was found that the prior expectations held by both leaders and other group members effect the formation of team norms, and that those leaders who held high prior expectations for team problem solving behaviours are able to compensate for a team's lower expectations. Taggar and Ellis' (2007) findings therefore suggest that leaders play a critical role in forming and shaping team norms. Taggar and Ellis (2007) further

asserted that effective team behaviours are improved through formal norms being established at the beginning of team interaction. If these norms are indeed influenced more by leaders than by the other members of the group, as suggested by Taggar and Ellis' (2007) findings and by Feldman (1984), and if unique information provided by experts is weighted more heavily than information provided by other members of the team, it can once again create a false sense of congruence in decision making. That is, while the team may agree to a particular decision, the level of importance placed on the opinions and wants of the 'leader' may be higher than those proposed by a group member ranking lower in the hierarchy, and may be discussed to a greater extent, thus leading to unique information that is perhaps as important or more important than that being considered being sidelined in discussions. The decisions made based on the discussions may, in such cases, be biased towards the overall opinions and wants of the team leader, rather than representing the opinions of the team members.

It should be pointed out that in group formation, team leaders are not necessarily those who are explicitly assigned the role, but are often assigned the role of leader through peer negotiation, spoken or unspoken (Seers, Petty, & Cashman, 1995). Leaders are often members of a group who are able to exert influence or authority over the team (Taggar & Ellis, 2007). Taking Kettles *et al.'s* (2009) findings into account, it may therefore be assumed that in a forensic or psychiatric multi-disciplinary team setting, the team leader is likely to be the individual perceived by the others to be most medically trained, e.g., psychiatrists, clinical psychologists, psychiatric nurses.

With regard to group decision making, it is important that all group members are happy with the hierarchical culture of the group. A successful hierarchical group culture should emphasise stability, rules, policy and regulations, while also accentuating teamwork, cohesiveness and participation (McComas, Tuite, Waks, & Sherman, 2007). A successful hierarchical group culture will therefore foster not only a more satisfied organisational culture, but also will increase individual group member's commitment (McComas *et al.*, 2007). The satisfaction of group members is important for optimal decision making in hospital and forensic cases making as it has been shown to improve performance and decrease the frustration that may otherwise be felt by members of a group (Zazzali, Alexander, Shortell, & Burns, 2007). In their research concerning group member's perceptions of procedural justice, Zazzali *et al.* (2007) pointed out that when members of a group are unsatisfied with a process, they are more likely to view it as biased and are less likely to accept the outcome. This may then lead to unsatisfied group members to lose interest and limit their participation in decision making discussions, an issue which, according to

Zazzali *et al.* (2007), has the potential to limit the ability of the group to make informed decisions as representative input is not achieved. To reiterate, when making judgements in violence risk assessment, not only should the uniqueness of information held by the group be made explicit and be taken into consideration, but so too should the diversity of the different group members in order to improve individual satisfaction and therefore group performance.

It may be argued that due to the relatively low volume of research surrounding descriptive research in the area of clinician judgement, that the most commonly over-emphasised cues used by clinicians have not been fully identified. This has major implications for the development of both actuarial scales and for improving clinician decision making alike. For example, in an assessment of the SAGE assessment survey, an actuarial tool measuring self-reported aggression, Downey and Zun (2007) found that while the overall scoring of the SAGE was unsuccessful at identifying those at risk of partaking in violent behaviour, certain factors in the tool were effective. While this would appear promising, the tool gives equal weight to all of the included factors, and thus the effect becomes annulled (Downy & Zun, 2007). Mills, Kroner, and Hemmati (2007) have further indicated that the inclusion of items that do not discriminate between recidivists and non-recidivists in risk assessment scales may indeed devaluate the risk assessment being conducted. This therefore suggests the need for the identification of the cues most often over-emphasised in terms of importance by both clinicians and risk assessment tools.

Ebogen's (2002) review of the descriptive research further highlighted issues regarding clinical reliability - for instance, the measurement of inter-clinician agreement on specific clinical judgements. It has been hypothesised that greater congruence in a decision made between clinicians will lead to greater reliability of that decision (Werner, Rose, & Yesavage, 1990). However, research by Quinsey and Ambtman (1979) found low levels of reliability between health professionals when assessing dangerousness. With regards to violence risk assessment, Werner, Rose, and Yesavage (1983) found similarly low levels of reliability in clinician decision making within the context of a psychiatric care facility when psychologists and psychiatrists were asked to make judgements of dangerousness. However, when the scores of judgements of dangerousness made by the psychologists and by the psychiatrists within the sample were combined as a single group, greater levels of congruence, thus reliability, were apparent in the violence risk assessments made. Thus it would appear that greater reliability in clinical decisions of violence risk assessment may be apparent within group decision making than would be the case with individual decision making.

Furthermore, research conducted by Skeem, Mulvey, Odgers, Schubert, Stowman, Gardner, and Lidz (2005) found that levels of violence prediction differed across mental health professionals pending on their professional role. In an examination of this disagreement between estimates of future violence occurring, Lidz, Mulvey, Apperston, Evanczuk, & Shea (1992) found that the key sources of this disagreement may be levels of clinician experience in addition to organisational variables. This finding would suggest that whist levels of accuracy increase with multiple decision makers being involved (Werner *et al.*, 1983), when the decision making group is of multiple disciplines/levels of experience, greater levels of disagreement within that group may exist (Lidz *et al.*, 1992). The level of congruence in clinical decisions made in groups is of utmost importance in terms of the reliability of that decision: it is believed that greater accuracy in violence risk assessments may be achieved in group decisions made with high levels of agreement (Elbogen, 2002).

As there are increasing requirements within violence risk assessment practice to make decisions in groups (Croskerry, 2005), it is of great importance that the implications of this alteration are understood. For Example, the types of heuristics and biases used by all decision makers unconsciously (Croskerry, 2005) are argued to be different for individual and group decision making. Thus, where the majority of past research in the field of clinical judgements of violence risk assessment has focused on individual decision makers, the findings of such research may not translate accurately with regards to group decisions in violence risk assessment.

Further to this increasing tendency of group decision making in violence risk assessment, there is little apparent standardisation of the demographics of the group decision makers. For example, in utilisation of the HCR-20 (Webster *et al.*, 1997) and other risk assessment/management tools, reports are gathered from multiple sources, e.g., social services, prison officers, clinical psychologists, parole officers, psychiatrists. While Andres Pueyo and Redondo Illescas (2007) point out that risk assessment is an individual process for the majority of mental health professionals and criminology experts involved in the task, through gathering data from multiple sources, as outlined above, and through different means (e.g., personal interviews, psychological and medical assessments, judicial and social records and the collateral recollection of information) and then discussing the initial individual judgements made with the multi-disciplinary team, the violence risk assessment process may then be considered, in part, as a multi-disciplinary process. The use of multi-disciplinary decision making groups with varying levels of expertise in clinical violence risk assessment has neglected to be the primary focus of the majority of research in the area. Whilst research, such as that by Werner *et al.* (1983), indeed investigated the differences between different types of clinical decision makers,

subsequent research in the area has not appeared to progress to such a level as to explore the possible social impact of multi-disciplinary group decision making in violence risk assessment, in comparison to single-disciplinary group or individual clinical decision makers.

In addition, Elbogen's (2002) review highlighted the need for greater research in the area of clinical decision making on the heuristics and biases used by clinicians whilst making violence risk assessments. Further to this, McNeil (1998) suggested that in order to improve clinician skill when making violence risk assessments, there must be at least an awareness of the possible cognitive biases that may be active with decision making.

The present article has summarised the key points relating to the state of violence risk assessment, informing the reader of the use of clinical judgement, actuarial tools, and structured professional guidelines when conducting a violence risk assessment. The controversy surrounding clinical judgement and the actuarial assessment of violence risk in particular has been discussed. In addition, the present article has highlighted the need for a greater focus on the ways in which risk assessment and management could be improved and informed through more targeted research investigating the actual process of violence risk assessment (e.g., Elbogen, 2002), and the need for research investigating the possible social effects of multi-disciplinary group decision making, and the impact of cognitive heuristics and biases on the accuracy and efficacy of violence risk assessments.

References

American Bar Association (1989). *Criminal justice mental health standards*. Washington, D. C.

Andres Pueyo, A., & Redondo Illescas, S. (2007). Dangerousness and violence risk assessment. *Papeles del Psicólogo*, 28(3), 157-173.

Ashley, M. (1922). Outcome of 1,000 cases paroled from the Middletown State Hospital. *State Hospital Quarterly*, 8, 64-70.

Berg, A., Bell, C. C., & Tupin, J. (2000). Clinician safety: Assessing and managing the violent patient. In C. Bell (Ed.), *Psychiatric aspects of violence: Issues in prevention and treatment*. San Francisco: Jossey-Bass.

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

Bowers, C. A., Pharmer, J. A., & Salas, E. (2000). When member homogeneity is needed in work teams: A meta analysis. *Small Group Research*, 31, 305-327.

Canter, D. V., & Zukauskiene, R. (2008). *Psychology and law: Bridging the gap*. Surrey: Ashgate Publishing Limited.

Carroll, A. (2007). Are violence risk assessment tools clinically useful? *Australian and New Zealand Journal of Psychiatry*, 41, 301-307.

Cooke, D. J., Michie, C., & Ryan, J. (2001). *Evaluating risk for violence: a preliminary study of the HCR-20, PCL-R and VRAG in a Scottish prison sample*. Scottish Prison Service Occasional Paper Series 5/2001.

Davidson, S. (1997). Risk assessment & management – A busy practitioner's perspective. *International Review of Psychiatry*, 9, 201-206.

Dawes, R. M., Faust, D., & Meehl, P. E. (1989). Clinical versus actuarial judgment. *Science*, 243, 1668-1674.

Dempster, R. J. (2003). Issues in the assessment, communication, and management of risk for violence. In W. T. O'Donohue and E. R. Levensky (Eds), *Handbook of forensic psychology: Resource for mental health and legal professionals* (pp.103-127). Sydney, Australia: Elsevier Academic Press.

Douglas, K. S., Ogloff, J. R. P., & Hart, S. D. (2003). Evaluation of a model of violence risk assessment among forensic psychiatric patients. *Psychiatric Services*, 54, 1372-1379.

Downey, L. D., & Zun, L. S. (2007). Violence prediction in the emergency department. *Violence: Recognition, Management and Prevention*, 9(1), 51-58.

Elbogen, E. B. (2002). The process of violence risk assessment: A review of descriptive research. *Aggression and Violent Behaviour*, 7, 591-604.

Elbogen, B. E., & Huss, M. T. (2000). The role of serotonin in violence and schizophrenia: Implications for risk assessment. *Journal of Psychiatry and Law*, 28, 19-48.

Feldman, D. C. (1984). The development and enforcement of group norms. *Academy of Management Review*, 9, 47-53.

Garfield, F. B., & Garfield, J. M. (2000). Clinical judgment and clinical practice guidelines. *International Journal of Technology Assessment in Health Care*, 16(4), 1050-1060.

Greenland, C. (1980). Psychiatry and the prediction of dangerousness. *Journal of Psychiatric Treatment and Evaluation*, 2, 205-222.

Hall, H. V. (1982). Dangerous predictions and the maligned forensic professional: Suggestions for detecting distortions of true basal violence. *Criminal Justice and Behavior*, 9, 3-12.

Harris, A., & Lurigio, A. J. (2007). Mental illness and violence: A brief review of the research and assessment strategies. *Aggression and Violent Behavior*, 12, 542-551.

Hart, S. D. (2000). The promise and peril of sex offender risk assessment. In R. Laws (Chair), *Structured professional guidelines for assessing risk in sexual offenders*. Symposium presented at the Annual Conference of the Association for the Treatment of Sexual Abusers, San Diego, California.

Hart, S. D., Michie, C., & Cooke, D. J. (2007). Precision of actuarial risk assessment instruments. *British Journal of Psychiatry*, 190(49), s60-s65.

Heilbrun, K. (1996). When is an "expert" an expert? *American Psychology Law Society*, 16(3), 5-6.

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

Hollinghead, A. B. (1996). The rank-order effect in group decision making. *Orginizational Behavior and Human Decision Processes*, 68(3), 181-193.

Homan, A. C., Van Knippenberg, D., Van Kleef, G. A., & De Dreu, C. K. W. (2007). Interacting dimensions of diversity: Cross-categorization and the functioning of diverse work groups. *Group Dynamics: Theory, Research, and Practice*, 11(2), 79-94.

Howard, P. (2009). *Improving the prediction of re-offending using the Offender Assessment System*. Ministry of Justice Research Summary 2/09.

Johnstone, L., & Cooke, D. J. (2008). *PRISM: Promoting Risk Intervention by Situational Management. Structured professional guidelines for assessing situational risk factors for*

violence in institutions. Burnaby, British Columbia, Canada: Simon Fraser University, Mental Health, Law, and Policy Institute.

Kettles, A., Gass, J., Addo, M., & Mckie, A. (2009). *Decision making about observation, engagement and risk assessment in IPCU's in Scotland*. Paper presented at Forging the Future: A Showcase of Innovation, Developments and Research in Scotland's Forensic Mental Health Services, Clydebank, November 2009.

Klassen, D., & O'Conner, W. (1987). Predicting violence in mental patients: Cross-validation of an actuarial scale. Paper presented at the Annual Meeting of the American Public Health Association.

Kozol, H. L., Boucher, R. J., & Garofalo, R. F. (1972). The diagnosis and treatment of dangerousness. *Crime in Delinquency*, 19, 371-392.

Kozol, H. L., Boucher, R. J., & Garofalo, R. F. (1973). Dangerousness: A reply to Monahan. *Crime and Delinquency*, 19, 554-555.

Kriegman, D. (2007). Affidavit in support of a Daubert challenge to the admissibility of scientific testimony to support Sexually Dangerous Person commitment under the Adam Walsh Act, Case 1:06-mc-PBS Documents 12-1 to 12-13, Filed 16/05/2007.

Lamont, S., & Brunero, S. (2009). Risk analysis: An integrated approach to the assessment and management of aggression/violence in mental health. *Journal of Psychiatric Intensive Care*, 5(1), 25-32.

Lidz, C., Mulvey, E., Apperson, L. J., Evanczuk, K., & Shea, S. (1992). Sources of disagreement among clinicians' assessments of dangerousness in a psychiatric emergency room. *International Journal of Law and Psychiatry*, 15, 237-250.

Litwack, T. R. (2001). Actuarial versus clinical assessments of dangerousness. *Psychology, Public Policy and Law*, 7(2), 409-443.

Litwack, T. R., Zapf, P. A., Groscup, J. L., & Hart, S. D. (2006). Violence risk assessments: Research, legal, and clinical considerations. In I. B. Weiner and A. K. Hess (Eds), *The Handbook of Forensic Psychology,* New York: Wiley and Sons, Inc.

McComas, K., Tuite, L. S., Waks, L., & Sherman, L. A. (2007). Predicting satisfaction and outcome acceptance with advisory committee meetings: The role of procedural justice. *Journal of Applied Social Psychology*, 37(5), 905-927.

McNeil, D. E. (1998). Empirically based clinical evaluation and management of the potentially violent patient. In P. M. Kleepsies (Ed.), *Emergencies in mental health practice: Evaluation and management* (pp.1-19). New York: Guilford Press.

Menzies, R. J., & 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). The dimensions of dangerousness: Evaluating the accuracy of psychometric predictions of violence among forensic patients. *Law and Human Behaviour*, 9, 49-70.

Monohan, J. (1984). The prediction of violent behaviour: Toward a second generation of theory and policy. *American Journal of Psychiatry*, 141, 458-471.

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

Quinsey, V. L. (1995). The prediction and explanation of criminal violence. *International Journal of Law and Psychiatry*, 18, 117-127.

Quinsey, V. L. & Ambtaman, R. (1979). Variables affecting psychiartists' and teachers' assessments of the dangerousness of mentally ill offenders. *Journal of Consulting and Clinical Psychology*, 47, 353-362.

Quinsey, V. L., & Cyr, M. (1986). Perceived dangerousness and treatability of offenders: the effects of internal versus external attributions of crime causality. *Journal of Interpersonal Violence*, 1, 458-471.

Quinsey, V. L., Harris, G. T., Rice, M. E., & Cormier, C. A. (1998). *Violent offenders: Apparising and managing risk* (4th Ed.). Washington, DC: American Psychological Association.

Rabinowitz, J., & Garelik-Wyler, R. (1999). Accuracy and confidence in clinical assessment of psychiatric inpatients risk of violence. *International Journal of Law and Psychiatry*, 22(1), 99-106.

Scott, C. L., & Resnick, P. J. (2006). Violence risk assessment in persons with mental illness. *Aggression and Violent Behaviour*, 11, 598-611.

Seers, A., Petty, M. M., & Cashman, J. F. (1995). Team-member exchange under team and traditional management: A naturally occurring quasiexperiment. *Group and Orginizational Mangement*, 20, 18-38.

Skeem, J. L., Mulvey, E. P., Odgers, C., Schubert, C., Stowman, S., Gardner, W., & Lidz, C. (2005). What do clinicians expect? Comparing envisioned and reported violence for male and female patients. *Journal of Counselling and Clinical Psychology*, 73(4), 599-609.

Stasser, G., & Titus, W. (1985). Pooling of unshared information in group decision making: Biased information sampling during discussion. *Journal of Personality and Social Psychology*, 48, 1467-1478.

Stasser, G., Stewart, D. D., & Wittenbaum, G. M. (1995). Expert roles and information exchange during discussion: The importance of knowing who knows what. *Journal of Experimental Social Psychology*, 31, 244-265.

Steadman, H. J., & Cocozza, J. J. (1974). *Careers of the criminally insane: Excessive social control of deviance*. Lexington, MA: Lexington Books.

Stewart, D. D., & Stasser, G. (1995). Expert role assignment and information sampling during collective recall and decision making. *Journal of Personality and Social Psychology*, 69, 619-628.

Taggar, S., & Ellis, R. (2007). The role of leaders in shaping formal team norms. *The Leadership Quarterly*, 18, 105-120.

Thornberry, T. P., & Jacoby, J. E. (1979). *The criminally insane: A community follow-up of mentally ill offenders*. Chicago: University of Chicago Press.

Van Knippenberg, D., De Dreu, C. K. W., & Homan, A. C. (2004). Work group diversity and group performance: An integrative model and research agenda. *Journal of Applied Psychology*, 98, 1008-1022.

Webber, S. S., & Donahue, L. M. (2001). The impact of highly and less job-related diversity on work group cohesion and performance: A meta-analysis. *Journal of Management*, 27, 141-162.

Webster, C. D., Douglas, K. S., Eaves, D., & Hart, S. D. (1997). *HCR-20. Assessing the risk of violence. Version 2*. Vancouver, Canada: Simon Fraser University and Forensic Psychiatric Servvices Commission of British Columbia.

Werner, P., Rose, T. L., & Yesavage, J. A. (1983). Reliability, accuracy, and decision-making strategy in clinical predictions of imminent dangerousness. *Journal of Consulting and Clinical Psychology*, 51, 815-825.

Werner, P., Rose, T. L., & Yesavage, J. A. (1990). Aspects of consensus in clinical predictions of imminent violence. *Journal of Clinical Psychology*, 46(4), 534-538.

Zazzali, J. L., Alexander, J. A., Shortell, S. M., & Burns, L. R. (2007). Organizational culture and physician satisfaction with dimensions of group practice. *Health Services Research*, 42(3 Pt 1), 1150-1176.

About the authors:

Jennifer Murray*
Jennifer Murray is a PhD student in the Department of Psychology, Glasgow Caledonian University. Her Thesis investigates the effects of attributional manipulations in clinical judgements of violence risk assessment, at various stages of the violence risk assessment process. The current paper is completed in part-fulfilment of Jennifer Murray's PhD thesis.

*Requests for reprints should be addressed to Jennifer Murray, Dept. of Psychology, Glasgow Caledonian University, 70 Cowcaddens Road, Glasgow, G4 0PP, Scotland, UK (email: Jennifer.murray@gcal.ac.uk)

Dr. Mary E. Thomson
Dr. Mary E. Thomson is a reader of Decision Science from the Department of Psychology, Glasgow Caledonian University, who has published in various books and journals, including Risk Analysis, Decision Support Systems, the International Journal of Forecasting, the European Journal of Operational Research and the Journal of Behavioral Decision Making.