A review of effective interventions for reducing aggression and violence

James McGuire*

Division of Clinical Psychology, School of Population, Community and Behavioural Sciences, University of Liverpool, Whelan Building, Liverpool L69 3GB, UK

This paper addresses the question of whether individual violence can be reduced in frequency or severity, if so to what extent and by which methods. It opens with a brief overview of the nature of personal violence and discussion of some key definitional and methodological problems. However, its principal focus is on the findings obtained from a series of meta-analytic reviews of structured programmes for adolescents and adults who have shown repeated aggression or been convicted of personal violence, drawing together the results of studies conducted in prison, probation, youth justice and allied services. Additional results are considered from a systematic review of studies of violence prevention among offenders with mental disorders. This incorporates the preliminary findings of a meta-analysis of controlled trials of psychosocial interventions with that population. Overall, it is concluded that there is sufficient evidence currently available to substantiate the claim that personal violence can be reduced by psychosocial interventions, but that much more research is required to delineate the parameters of effectiveness in this context. Proposals are made for future investigations with reference to the theoretical understanding of causal relationships and the design of experimental trials.

Keywords: aggression; violence; recidivism; interventions; programmes

1. Objectives
The principal objectives of this paper are to survey research on the outcomes of interventions designed to reduce personal violence and to summarize what has emerged from that work with a view to identifying the most effective approaches to the problem that have been discovered to date. The emphasis will be primarily on aggression and violent behaviour among adolescents and adults, and although work with younger children will be discussed it is not the central focus of this paper.

To facilitate the principal task, three initial but subsidiary objectives will be briefly addressed. They are as follows: first, to consider some issues that arise in defining violence, in order to obtain a clearer appraisal of the subject under discussion; second, to survey the nature of personal violence as a social and public health problem; and third, to discuss some difficulties that arise when conducting research in this area, with particular reference to the evaluation of interventions. A fourth objective, addressed later in this paper, is to forward an integrative perspective on the factors that influence the occurrence of violence acts, in a probabilistic causal model.

There is a useful distinction that is often made in studying interventions designed to reduce recurrent problem behaviours such as criminal conduct in general or violent offending, in particular. This is a classification forwarded by Guerra et al. (1994) between primary, secondary and tertiary preventions of antisocial activity in adolescence.

The first of these refers to general-population or community-level initiatives designed to alleviate a social problem. In criminology, this may consist of situational prevention, e.g. 'target hardening', increased security, improvements in street lighting or other environmental modifications, installation of closed-circuit television, neighbourhood watch or other efforts to make crimes less likely to occur. Alternatively it may consist of investment in additional resources, e.g. improved family welfare, education or other social provision that will lower the socioeconomic deprivation associated with some types of crime (Farrington & Coid 2003).

The second denotes any type of intervention that 'targets individuals who show preclinical manifestations of some types of problem, whether it be physical, psychological or social' (Fields & McNamara 2003, p. 66). This entails work with designated 'at-risk' groups, such as children truanting from school, or involved in bullying or other aggression that is not strictly classed as illegal (owing to the perpetrator's age or other characteristics) (Goldstein 2002).

The second denotes any type of intervention that 'targets individuals who show preclinical manifestations of some types of problem, whether it be physical, psychological or social' (Fields & McNamara 2003, p. 66). This entails work with designated 'at-risk' groups, such as children truanting from school, or involved in bullying or other aggression that is not strictly classed as illegal (owing to the perpetrator's age or other characteristics) (Goldstein 2002).

The second denotes any type of intervention that 'targets individuals who show preclinical manifestations of some types of problem, whether it be physical, psychological or social' (Fields & McNamara 2003, p. 66). This entails work with designated 'at-risk' groups, such as children truanting from school, or involved in bullying or other aggression that is not strictly classed as illegal (owing to the perpetrator's age or other characteristics) (Goldstein 2002).

The third level, tertiary prevention, describes work undertaken with adjudicated offenders, those already convicted by the courts (usually more than once), that is intended to reduce their subsequent rates of criminal recidivism (Gendreau & Andrews 1990). While some references will be made to the first two of these classifications in what follows, the primary
focus of this paper is on interventions in the criminal justice system, youth justice, probation, prison and related agencies.

2. THE NATURE OF VIOLENCE

Violence is not a unitary phenomenon: it takes many different forms. It is important to identify its principal manifestations and if possible to distinguish various types within them.

(a) Definitional problems

The central focus of this paper is on what is generally termed personal violence, committed by one individual against another, in a context in which such actions are specifically proscribed by societal norms and usually codified in formal law. Under the legal framework of England and Wales, for example, numerous statutes, perhaps most notably the Offences Against the Person Act (1861), define the nature of actions that are considered to constitute different forms of violent crime (e.g. assault occasioning actual bodily harm, grievous bodily harm, malicious wounding, threats to kill). While this might appear straightforward, the task of describing and classifying violent behaviour in legal terms is a complex and challenging one (Carter & Harrison 1991). English law discerns numerous other discrete types of violence, for example, harassment, racially or religiously aggravated assault, public disorder, sexual assault, infanticide, homicide and complicity in another’s suicide (Cook et al. 2006).

The associated complexities notwithstanding, one way to define violence is to do so in terms of what is described as such in the criminal law. Much research in social science adopts that approach.

Definitions of violence have been described as varying along six dimensions (Jackson & Brownstein 2004). They are: (i) the level of action of the behaviour (individual, interpersonal, collective), (ii) the nature and degree of force, (iii) the outcome including extent of injury, (iv) the type of injury (e.g. physical harm, emotional degradation, interpersonal dominance), (v) the nature and significance of the target(s), and (vi) whether or not the actions were intentional. Given the multiple possibilities so generated, there is as yet no agreed benchmark regarding how violence should be conceptualized.

Thus, violent offences can be defined restrictively in terms of physical harm, possibly (as done in the Denver Youth Survey) subdividing them to denote ‘those acts in which someone was hurt or injured, but perhaps only in a minor way. Thus, hitting, getting into fights and so on are included. Serious violent offences, however, include only acts that resulted in serious injury (requiring medical treatment—cut, bleeding, unconscious, etc.) or in which a weapon was used’ (Thornberry et al. 1995, p. 224). In the field of developmental psychopathology by contrast, violence is generally conceptualized more broadly, to encompass acts committed by children below the age of criminal responsibility and to study continuities in behavioural patterns across successive maturational periods. For example, in forwarding a model of the development of antisocial behaviour (AB) and conduct problems, Dodge & Pettit (2003) defined the target variables as including ‘recurrent problem behaviours that lead to injury to others or arrest’ (2002, p. 350), so referring not only to violent crimes in the sense delineated above, but also ‘...verbal assault, vandalism, delinquency, destruction of another’s social standing (as in relational aggression), and physical abuse of one’s offspring’ (2003, p. 350). These actions have interpersonal targets and are distinguished from self-destructive behaviours such as substance abuse, suicide and internalizing problems. Given the observed continuities in these behaviours across developmental stages, in the present paper this broader definition will be employed, to include research on interventions designed to reduce any repeated manifestations of interpersonally harmful behaviour.

(b) Spatial and temporal variations

Personal violence as defined in criminal law typically constitutes a smaller fraction of all crimes than offences against property (such as theft, burglary or criminal damage). For example, in England and Wales in 2006/2007, it formed 22% of recorded crime (Home Office 2007). Although officially recorded violent crime doubled between 1998/1999 and 2006/2007, there is evidence that the underlying rate of total and of violent crime has been generally declining. Since 1995 incidents of wounding have fallen by 37%, assaults with minor injury by 58% and assaults resulting in no injury by 36%. Nevertheless, according to British Crime Survey estimates, there were still nearly 2.5 million incidents of violence against adults in 2006/2007 (Jansson et al. 2007).

The amounts and levels of seriousness of personal violence show considerable international variations (Bureau of Justice Statistics 2003). Table 1 shows selected extracts from data collected by the United Nations Office on Drugs and Crime (2008), showing recorded rates of intentional homicide, serious assault and rape in 15 UN member countries for the year 2002 (the latest year for which comparative figures are available). Those listed are designed to illustrate the sizeable disparities in officially recorded rates. Victim surveys where available tend to show lower rates of reporting for violence than for other types of crimes (Van Kesteren et al. 2000; Del Frate 2003; Naudé et al. 2006).

A separate series of surveys is reported by the World Health Organization (WHO; Krug et al. 2002). Focusing on deaths caused by violence, as a broad initial categorization, this first of all distinguishes deaths due to suicide, military conflicts and homicide, respectively. Surveying data on a global scale, Krug et al. estimated that in the year 2000 there were approximately 1.66 million deaths due to violence. Of this total, just under half (815 000) consisted of suicides, roughly one-fifth (310 000) were war related and one-third (510 000) were homicides. The last figure corresponds to one approximately every 60 s.

As table 1 shows, however, the homicide rate is only the ‘tip of the iceberg’ as far as crimes of violence are concerned. Non-fatal assaults are far more numerous (those shown are only the most serious). The WHO report (Krug et al. 2002) also cites figures on domestic violence.
many studies indicate (as is found with most other types of offence) that a relatively small proportion of the population is responsible for a comparatively large proportion of recorded violent crimes (Surgeon General 2001). For example, in the Rochester Youth Development Study, it was found that 15% of the sample accounted for 75% of the violent offences (Thornberry et al. 2003). According to Tremblay (2003), the proportion of children identified in longitudinal studies as exhibiting chronic physical aggression was approximately 5%. According to Blair et al. (2005), the proportion of adults repeatedly engaged in violence as a function of psychopathy was also approximately 5%.

Thus there are measurable individual differences in aggressiveness, which have been conceptualized as ‘...a relatively persistent readiness to become aggressive in a variety of different situations’ (Berkowitz 1993, p. 21). Concerning individuals likely to act frequently or repeatedly in this way, there is evidence of relative stability in patterns of aggressiveness between infancy, middle childhood, adolescence and adulthood, comparable with that shown for general intellectual functioning. For example, the presence of aggressive classroom behaviour in the early school years has been shown to be a good predictor of delinquency in adolescence (Spivack & Cianci 1987). Across a more limited time scale of 5 or 6 years, studies have shown that aggressive behaviour in middle childhood is strongly predictive of conduct problems during the teenage years (Loeber & Stouthamer-Loeber 1988; Farrington & West 1993). A recent study has shown moderate stability in the assessments of psychopathy across an 11-year period, between ages 13 and 24 (Lynam et al. 2007).

Data of these kinds have been consolidated in two independent analyses of correlations over time. Olweus (1979, 1988) reported a review of 16 longitudinal studies examining the levels of consistency in aggressive behaviour over periods ranging from 1 to 21 years. The dependent variables in the studies were nominations or ratings of aggressiveness provided by peers, teachers or other observers. From these findings, Olweus extracted

### Table 1. International variations in selected crimes of personal violence. (Police statistics, year 2002, rates per 100 000 population. Source. United Nations Office on Drugs & Crime.)

<table>
<thead>
<tr>
<th></th>
<th>Intentional Homicide</th>
<th>Major Assault</th>
<th>Rape</th>
</tr>
</thead>
<tbody>
<tr>
<td>South Africa</td>
<td>47.53</td>
<td>587.32</td>
<td>South Africa 115.61</td>
</tr>
<tr>
<td>El Salvador</td>
<td>31.54</td>
<td>343.39</td>
<td>Canada 77.64</td>
</tr>
<tr>
<td>Mexico</td>
<td>13.04</td>
<td>310.14</td>
<td>United States 32.99</td>
</tr>
<tr>
<td>Belarus</td>
<td>9.96</td>
<td>186.68</td>
<td>United Kingdoma 22.62</td>
</tr>
<tr>
<td>Argentina</td>
<td>9.47</td>
<td>152.22</td>
<td>Mexico 14.26</td>
</tr>
<tr>
<td>Lithuania</td>
<td>8.45</td>
<td>137.28</td>
<td>El Salvador 13.12</td>
</tr>
<tr>
<td>Uruguay</td>
<td>6.46</td>
<td>70.78</td>
<td>Denmark 9.30</td>
</tr>
<tr>
<td>Albania</td>
<td>5.68</td>
<td>43.70</td>
<td>Uruguay 9.02</td>
</tr>
<tr>
<td>United States</td>
<td>5.62</td>
<td>39.03</td>
<td>Belarus 8.59</td>
</tr>
<tr>
<td>United Kingdoma</td>
<td>2.03</td>
<td>25.38</td>
<td>Argentina 8.32</td>
</tr>
<tr>
<td>Poland</td>
<td>1.87</td>
<td>23.33</td>
<td>Poland 6.13</td>
</tr>
<tr>
<td>Canada</td>
<td>1.67</td>
<td>13.97</td>
<td>Lithuania 5.42</td>
</tr>
<tr>
<td>Tunisia</td>
<td>1.22</td>
<td>13.14</td>
<td>Morocco 3.42</td>
</tr>
<tr>
<td>Denmark</td>
<td>1.04</td>
<td>8.56</td>
<td>Tunisia 3.13</td>
</tr>
<tr>
<td>Morocco</td>
<td>0.48</td>
<td>3.71</td>
<td>Albania 1.43</td>
</tr>
</tbody>
</table>

a Per capita values for England and Wales were considered a proxy for the entire UK.

### (c) Individual continuity

With reference to the more regularly occurring types of aggression and violence as conventionally defined,
a total of 24 correlation coefficients and plotted their interrelationships on a regression line. While unsurprisingly the correlations decreased with increasing time intervals, there was nevertheless a striking degree of consistency. For example, mean correlations across 12 and 24 months were 0.76 and 0.69, respectively, but fell to 0.36 in one 21-year follow-up. In a subsequent review, Zumkley (1994) analysed an additional 10 studies, confirming the pattern found by Olweus. The levels of continuity discovered in these studies may be a function of the subgroup of those engaged in AB, who can be depicted as manifesting life-course-persistent delinquent conduct, as hypothesized by Moffitt (1993, 2003).

(d) Expressive and instrumental aggression
Attempts to identify subtypes of aggression are particularly important with reference to the small proportion of individuals likely to manifest persistently high levels of it over a prolonged segment of the life course (Hodgins 2007). Researchers and clinicians have found it useful to make a distinction between the following two types of aggression: expressive and instrumental. In the former, also variously called reactive, angry, emotional, hostile or impulsive aggression, harm to a victim decreases an unpleasant internal feeling state in the aggressor, possibly through the reduction of physiological arousal or tension (Berkowitz 1993; Blackburn 1993). In the latter, threats or injury facilitate achievement of non-injurious goals as, for example, in robbery; violence works proactively as a means to an end rather than an end in itself. While some researchers have expressed reservations over the rigour with which this distinction can be maintained, given that many aggressive acts have mixed motives (Bushman & Anderson 2001; Anderson & Bushman 2002), there is both psychometric and neuropsychological evidence that supports its meaningfulness (Blair et al. 2005). In addition, the distinction appears vital when allocating individuals to intervention programmes, given the importance of linking interventions to a functional understanding of violence motivation.

3. CHALLENGES OF VIOLENCE INTERVENTION RESEARCH
There are several major obstacles to pursuing a systematic scientific study of violent behaviour, and additional difficulties arise in research designed to evaluate the impact of tertiary interventions. First, as already described, defining violence is a precarious and unsatisfactory process. There is consequently a fundamental and recurrent problem of recording and measurement in this field. Studying aggression or other behaviour labelled as ‘antisocial’ in children, it is recognized that many acts though superficially similar are functionally very dissimilar (Tremblay 2000). The experience of an interpersonal event such as an assault and how it is described, respectively, by protagonists and observers is a product of both individual subjective perceptions and wider socially constructed interpretative frameworks. Second, notwithstanding the data presented above on the numbers of violent crimes, relative to other human activities violence is a comparatively rare event. Even the most frequently aggressive individual is unlikely to act violently in every encounter. That serious and violent offending manifests an ‘episodic or intermittent nature...has important implications for research’ (Huizinga et al. 2003, p. 55). The low base rate of assault can make it very difficult to isolate a group of individuals who can be reliably identified as violence-prone: most individuals who commit such offences are criminally ‘versatile’. When evaluating the impact of interventions, the low frequency of violent reoffences may result in studies being statistically underpowered.

Third, most assaults are a product of an intricate sequence of events and processes of separate but interconnected types (neurobiological, hormonal, cognitive, attitudinal, experiential, interactional). Hostile exchanges resulting in violence are influenced by the interplay of dispositional and situational factors. Those in turn are a function of participants’ temperaments, developmental histories, socialization experiences, interpersonal skills, attitudes and self-concepts (Toch 1969; Tremblay 2000). Identifying causal pathways and developing integrative explanatory models pose major challenges.

Fourth, commensurate with the level of complexity of violence causation, when individuals who are frequently aggressive are assessed by criminal justice or other practitioners, they are usually found to have multiple criminogenic needs. Interventions devised to address several problems in combination have been termed multi-modal (Lipsey 1995). However, their implementation leaves a significant residual problem of disentangling the most likely ‘active ingredients’ and of developing appropriate methods for maximizing treatment impact. Research projects in which such components are dismantled and evaluated separately remain relatively exceptional, leaving many issues concerning effectiveness unresolved.

Fifth, in outcome research there are often competing demands between practical need and service delivery on the one hand, and rigorous evaluation for the purposes of hypothesis testing on the other. Allocation to different levels or types of services in social welfare and criminal justice is customarily in the hands of decision-makers (sentencers, case managers) whose priorities are in the realms of public protection or of meeting client needs. Typically, evaluation projects are unlikely to attain the standards of good experimental designs, most importantly in ensuring the equivalence of experimental and control samples with respect to key variables that may be conflated with measured outcomes. An additional difficulty arises from the recurrently high levels of attrition found among offender samples. These factors often reduce the methodological quality of evaluations with important consequences for hypothesis testing regarding treatment effects.

4. REDUCING CRIMINAL RECIDIVISM
The feasibility of implementing effective tertiary prevention in criminal justice has been the subject of sometimes fierce debate for several decades. The associated controversies have been summarized
elsewhere (e.g. Gaes 1998; McGuire 2004; Andrews & Bonta 2006). Whereas until approximately the late 1990s there was a widely held assumption that persistent offending behaviour was not susceptible to change, the growing volume of positive research findings eventually achieved a sufficient critical mass to engender the pursuit of a variant of evidence-based practice in criminal justice services (Raynor 2004). The background evidence relevant to this will first be briefly summarized before turning attention to the findings more closely related to aggression and violence.

(a) General findings
The usage of meta-analysis has had a considerable impact on criminology with particular reference to offender treatment (Wilson 2001), and between 1985 and mid-2007 70 meta-analyses were published or presented in this and adjacent fields. These reviews are listed in table 2, which shows authors and dates of publication, the number of effect-size tests (k) conducted within each meta-analysis and the mean effect size obtained either across all studies reviewed or from specified subsets within them. The table includes basic information from meta-analytic reviews of interventions with adjudicated offenders or in some instances pre-delinquents (e.g. those exhibiting AB in school or home settings). Reviews are tabulated chronologically by the year of publication and in alphabetic order of first-named authors within each year.

This list excludes reviews of (i) evaluations of changes in penal policy or practice such as the introduction of curfews, drug courts or new sentencing guidelines, (ii) studies of psychological therapy conducted exclusively with children outside the age range of criminal responsibility and (iii) reviews of anger management and allied interventions to be discussed separately below.

The number of tests as listed in column 3 may not correspond to the number of studies that were subsumed in a review. Rather, it is the number of effect sizes used to compute a mean effect relevant to the primary focus of the review. Where more than one effect size is given in column 4, this corresponds to major categories reported in a review. Not all reviews were designed to test hypotheses about whether treatment 'worked'; several are focused on the importance of moderator variables (age, gender, ethnicity, treatment integrity); some are evaluations of interventions for specific types of offence (violent, sexual, drink-driving, substance abuse); and others of the impact of interventions conducted in different settings (prison, community, school, family).

Most effect sizes reported are standardized mean difference (d or g) or correlation coefficients (r or g). In these cases a + sign indicates that the outcome favoured the experimental sample (i.e. its level of recidivism was lower than that of controls). However, where odds ratios are reported, the direction varies according to the method of analysis used in the denoted review. Where column 4 reads na (not available/not applicable) this is where, for varying reasons, no average effect size was computed, or alternatively the overall result is not a comparison between treated and untreated groups, but a mean correlation between either an independent or moderator variable and recidivism outcomes.

As can be seen from the table almost all effect sizes are positive, and while some are close to zero the majority are in the small or moderate range, following the conventions proposed by Cohen (1992). The only negative mean effect sizes reported to date are those obtained from criminal sanctions/deterrence or treatment of offenders classified as psychopaths. One outcome of the generally encouraging pattern regarding psychosocial intervention has been the development and dissemination of structured programmes (Andrews 2001; McGuire 2001): pre-planned and reproducible sequences of focused activities; typically delivered in group formats; and supported by a manual and other accompanying materials.

(b) Limitations of studies and reviews
A degree of caution is warranted when scrutinizing the output from a large array of meta-analytic reviews such as that shown in table 2. Some researchers have been sceptical about the use of meta-analysis on a variety of grounds. First, if the quality of the original research is poor, regretfully a not infrequent occurrence, it will be neither feasible nor permissible to draw any firm conclusions even from the most carefully conducted review of it. Second, and as noted above, given the circumstances in which most research of this kind takes place, the design of some evaluation studies is very weak. It can be difficult to use random allocation to experimental and comparison samples, and the members of these groups are often not well matched, as researchers may have no control over who is placed under what conditions. As this activity often takes place in the daily routines of the criminal justice system, neatly designed experimental trials are difficult to carry out, and most evaluations are of a less robust nature (Lipsey 1999). Third, follow-up periods have often been very short: six to nine months is not uncommon. However, there are also studies with 12- and 24-month follow-up, and a proportion where data have been collected for five years or more (e.g. a mean of 46 months in the review by Hanson et al. 2002). Fourth, sample sizes in some studies are small and if there is a further loss due to attrition, or attempts are made to subdivide samples for particular analyses, it may be difficult to draw clear conclusions. Fifth, despite there being hundreds of primary studies in this field, when examined more closely the number in any given category can be disappointingly small (Lösel 2001). Thus when attempting to review studies, it can be difficult to draw anything other than the most 'broad-brush' conclusions (Lipsey 1995). Sixth, it has been claimed that positive outcomes are merely a product of self-selection effects: if offenders participating in treatment are observed to change, it is primarily because they were motivated to do so, and would do well anyway (Simon 1998; but see Hollin 2006). Finally, an awkward interpretative problem arises due to publication bias. If research studies with non-significant findings are less likely to be submitted and/or published, those that are publicly available may be unrepresentative of the research actually done.
Table 2. Summary information from 70 meta-analyses of offender treatment-outcome studies 1985–2007.

<table>
<thead>
<tr>
<th>author(s) and year of publication</th>
<th>focus of review</th>
<th>number of effect-size tests (k)</th>
<th>mean effect size(s) reported</th>
</tr>
</thead>
<tbody>
<tr>
<td>Garrett (1985)</td>
<td>young offenders in residential placements</td>
<td>121</td>
<td>+0.18</td>
</tr>
<tr>
<td>Gensheimer et al. (1986)</td>
<td>diversion schemes for young offenders</td>
<td>31</td>
<td>+0.26</td>
</tr>
<tr>
<td>Mayer et al. (1986)</td>
<td>social learning based interventions with youth</td>
<td>17</td>
<td>+0.33</td>
</tr>
<tr>
<td>Gottschalk et al. (1987a)</td>
<td>community-based interventions with youth</td>
<td>61</td>
<td>+0.22</td>
</tr>
<tr>
<td>Gottschalk et al. (1987b)</td>
<td>behavioural interventions with youth</td>
<td>14</td>
<td>+0.25</td>
</tr>
<tr>
<td>Lösel &amp; Koferl (1989)</td>
<td>socio-therapeutic prison regimes in Germany</td>
<td>16</td>
<td>+0.12</td>
</tr>
<tr>
<td>Whitehead &amp; Lab (1989)</td>
<td>young offenders: general</td>
<td>50</td>
<td></td>
</tr>
<tr>
<td>Andrews et al. (1990)</td>
<td>testing model of 'human service principles'</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Izzo &amp; Ross (1990)</td>
<td>cognitive versus non-cognitive interventions</td>
<td>46</td>
<td>ratio of mean effect sizes = 2.5 : 1</td>
</tr>
<tr>
<td>Lipsey (1992)</td>
<td>offenders aged 12–21</td>
<td>397</td>
<td>+0.10</td>
</tr>
<tr>
<td>Hall (1995)</td>
<td>sexual offending</td>
<td>12</td>
<td>+0.12</td>
</tr>
<tr>
<td>Wells-Parker et al. (1995)</td>
<td>drink-driving offences</td>
<td>215</td>
<td>8–9% reduction</td>
</tr>
<tr>
<td>Gendreau &amp; Goggin (1996)</td>
<td>deterrence and intermediate punishment</td>
<td>138</td>
<td>0.00</td>
</tr>
<tr>
<td>Cleland et al. (1997)</td>
<td>impact of age as moderator variable</td>
<td>659</td>
<td>n.a.</td>
</tr>
<tr>
<td>Redondo et al. (1997)</td>
<td>European structured programmes</td>
<td>57</td>
<td>+0.12</td>
</tr>
<tr>
<td>Lipsey &amp; Wilson (1998)</td>
<td>serious violent and sexual offending by youth</td>
<td>24</td>
<td>+0.10</td>
</tr>
<tr>
<td>Marsch (1998)</td>
<td>methadone maintenance for opiate-dependence</td>
<td>24</td>
<td>+0.14</td>
</tr>
<tr>
<td>Alexander (1999)</td>
<td>sexual offending</td>
<td>79</td>
<td>+0.10</td>
</tr>
<tr>
<td>Dowden &amp; Andrews (1999a)</td>
<td>programmes for women offenders</td>
<td>24</td>
<td>n.a.</td>
</tr>
<tr>
<td>Dowden &amp; Andrews (1999b)</td>
<td>young offenders: general</td>
<td>229</td>
<td>+0.09</td>
</tr>
<tr>
<td>Gallagher et al. (1999)</td>
<td>sexual offending</td>
<td>25</td>
<td></td>
</tr>
<tr>
<td>Polizzi et al. (1999)</td>
<td>sexual offending</td>
<td>13</td>
<td>n.a.</td>
</tr>
<tr>
<td>Redondo et al. (1999)</td>
<td>European structured programmes</td>
<td>32</td>
<td>+0.12</td>
</tr>
<tr>
<td>Dowden &amp; Andrews (2000)</td>
<td>interventions for violent offenders</td>
<td>52</td>
<td>+0.07</td>
</tr>
<tr>
<td>Egg et al. (2000)</td>
<td>treatment programmes in Germany</td>
<td>25</td>
<td>r = +1.12, OR = 1.9</td>
</tr>
<tr>
<td>Petrosino et al. (2000)</td>
<td>scared straight programmes</td>
<td>9</td>
<td>−0.01</td>
</tr>
<tr>
<td>Prendergast et al. (2000)</td>
<td>treatment of drug dependence</td>
<td>28</td>
<td>drug use +0.29</td>
</tr>
<tr>
<td>Wilson et al. (2000)</td>
<td>(treatment-comparison studies only)</td>
<td>17</td>
<td>crime +0.17</td>
</tr>
<tr>
<td>Wilson &amp; Lipsey (2000)</td>
<td>educational and vocational programmes, adults</td>
<td>53</td>
<td>OR = 1.52</td>
</tr>
<tr>
<td>Wilson &amp; Lipsey (2001)</td>
<td>wilderness challenge programmes</td>
<td>22</td>
<td>+0.18</td>
</tr>
<tr>
<td>Gendreau et al. (2001)</td>
<td>intermediate punishment</td>
<td>140</td>
<td>0.00</td>
</tr>
<tr>
<td>Latimer (2001)</td>
<td>family treatment</td>
<td>50</td>
<td>+0.15</td>
</tr>
<tr>
<td>Lipsey et al. (2001)</td>
<td>cognitive-behavioural interventions</td>
<td>14</td>
<td>OR = 0.66</td>
</tr>
<tr>
<td>MacKenzie et al. (2001)</td>
<td>correctional boot camps</td>
<td>44</td>
<td>OR = 1.02</td>
</tr>
<tr>
<td>Wilson et al. (2001)</td>
<td>school-based interventions</td>
<td>40</td>
<td>d = +0.04</td>
</tr>
<tr>
<td>Hanson et al. (2002)</td>
<td>sexual offending</td>
<td>43</td>
<td>OR = 0.81</td>
</tr>
<tr>
<td>Lipton et al. (2002a)</td>
<td>therapeutic communities</td>
<td>35</td>
<td>+0.14</td>
</tr>
<tr>
<td>Lipton et al. (2002b)</td>
<td>cognitive-behavioural interventions</td>
<td>68</td>
<td>+0.12</td>
</tr>
<tr>
<td>Prendergast et al. (2002)</td>
<td>programme factors in treating drug dependence</td>
<td>78</td>
<td>drug use g +0.33</td>
</tr>
<tr>
<td>Redondo et al. (2002)</td>
<td>European structured programmes</td>
<td>23</td>
<td>crime g +0.13</td>
</tr>
</tbody>
</table>

(Continued.)
Drawing conclusions from published work alone may give a distorted picture. Many of these factors can be corrected or taken into account in properly conducted meta-analysis. Studies with larger samples can be given more weight, and well- and poorly designed experiments can be evaluated separately to check whether they demonstrate broadly similar effects. Although publication bias cannot be eradicated, it can be minimized by making every possible effort to locate unpublished studies; or by computing the fail-safe or file-drawer n, the number of unpublished studies with zero or negative effect sizes that would be needed to discount or overturn an observed positive effect.

<table>
<thead>
<tr>
<th>author(s) and year of publication</th>
<th>focus of review</th>
<th>number of effect-size tests (k)</th>
<th>mean effect size(s) reported</th>
</tr>
</thead>
<tbody>
<tr>
<td>Woollenden et al. (2002)</td>
<td>family-based interventions</td>
<td>5</td>
<td>OR = 0.66</td>
</tr>
<tr>
<td>Andrews &amp; Bonta (2006)</td>
<td>restorative justice</td>
<td>67</td>
<td>+ 0.07</td>
</tr>
<tr>
<td>Dowden &amp; Andrews (2003)</td>
<td>family-based interventions</td>
<td>53</td>
<td>+ 0.21</td>
</tr>
<tr>
<td>Dowden et al. (2003)</td>
<td>effectiveness of relapse prevention</td>
<td>40</td>
<td>+ 0.15</td>
</tr>
<tr>
<td>Farrington &amp; Welsh (2003)</td>
<td>family-based interventions</td>
<td>135</td>
<td>post-test d + 0.38</td>
</tr>
<tr>
<td>Lösel &amp; Beelman (2003)</td>
<td>child skills training</td>
<td></td>
<td>follow-up d + 0.28</td>
</tr>
<tr>
<td>Wilson et al. (2003a)</td>
<td>school-based intervention programmes</td>
<td>522</td>
<td>+ 0.25</td>
</tr>
<tr>
<td>Wilson et al. (2003b)</td>
<td>impact of ethnicity as moderator</td>
<td>305</td>
<td>n.a.</td>
</tr>
<tr>
<td>Nugent et al. (2004)</td>
<td>victim-offender mediation</td>
<td>15</td>
<td>OR = 0.70</td>
</tr>
<tr>
<td>Andrews &amp; Dowden (2005)</td>
<td>programme and treatment integrity as moderator</td>
<td>273</td>
<td>n.a.</td>
</tr>
<tr>
<td>Landenberger &amp; Lipsey (2005)</td>
<td>treatment factors, cog-behavioural programmes</td>
<td>58</td>
<td>OR = 1.53</td>
</tr>
<tr>
<td>Latimer et al. (2005)</td>
<td>restorative justice</td>
<td>32</td>
<td>+ 0.07</td>
</tr>
<tr>
<td>Lösel &amp; Schmucker (2005)</td>
<td>sexual offending</td>
<td>80</td>
<td>+ 0.29</td>
</tr>
<tr>
<td>Visher et al. (2005)</td>
<td>employment programmes (community)</td>
<td>10</td>
<td>+ 0.03</td>
</tr>
<tr>
<td>Wilson et al. (2005a)</td>
<td>cognitive-behavioural group programmes</td>
<td>74</td>
<td>range: + 0.16 to + 0.49</td>
</tr>
<tr>
<td>Wilson et al. (2005b)</td>
<td>correctional boot camps</td>
<td>43</td>
<td>OR = 1.02</td>
</tr>
<tr>
<td>Andrews &amp; Dowden (2006)</td>
<td>‘risk-needs’ principles of case classification</td>
<td>374</td>
<td>significant support</td>
</tr>
<tr>
<td>Bradshaw et al. (2006)</td>
<td>victim offender mediation (juveniles)</td>
<td>15</td>
<td>+ 0.34</td>
</tr>
<tr>
<td>French &amp; Gendreau (2006)</td>
<td>reducing prison misconducts</td>
<td>104</td>
<td>+ 0.14</td>
</tr>
<tr>
<td>Holloway et al. (2006)</td>
<td>impact of drug treatment on criminal recidivism</td>
<td>28</td>
<td>OR 1.41–1.56</td>
</tr>
<tr>
<td>McCart et al. (2006)</td>
<td>relative effects of behavioural parent training (BPT) and cognitive-behavioural therapy (CBT)</td>
<td>BPT 32 CBT 45</td>
<td>+ 0.47 + 0.35</td>
</tr>
<tr>
<td>Mitchell et al. (2006)</td>
<td>incarceration-based drug treatment</td>
<td>66</td>
<td>OR = 1.37</td>
</tr>
<tr>
<td>Tong &amp; Farrington (2006)</td>
<td>Reasoning and Rehabilitation programme</td>
<td>25</td>
<td>OR = 1.16</td>
</tr>
<tr>
<td>Garrido &amp; Morales (2007)</td>
<td>institutionally-based interventions, violent youth</td>
<td>recidivism 30 serious offence 15</td>
<td>OR = 1.13 OR = 1.35</td>
</tr>
<tr>
<td>Tanasichuk &amp; Wormith (2007)</td>
<td>comparison between treated and untreated psychopaths</td>
<td>3</td>
<td>criminality − 0.10 violence + 0.03</td>
</tr>
</tbody>
</table>

5. VIOLENCE META-ANALYSES

Of particular interest for present purposes are those reviews that have focused specifically on interventions designed to reduce aggression and violence. With reference to young offenders, this area has been subject to narrative review (Fields & McNamara 2003) and findings of several meta-analyses concerning the impact of residential treatment on young offenders’ general delinquency have also been collated (Grietens & Hellinckx 2004). To the present author’s knowledge, however, no previous review has synthesized findings across available meta-analyses focused on outcomes of interventions designed to reduce aggression and violence among both adolescents and adults.
Most intervention studies, and consequently many meta-analyses, describe work with samples of offenders with varied criminal histories, a proportion of whom are likely to have committed assaults, robberies or other violent crimes in the midst of a pattern of ‘generalist’ offending. Most individuals with multiple convictions for violent offending have committed other types of offence in addition (Surgeon General 2001). Unfortunately, many studies give insufficient details of the criminal backgrounds of treatment participants, and the number reporting the differential impact of interventions on violent as opposed to other types of reconviction is not large. Where such data have been assembled, meta-analysis has revealed a trend towards larger effect sizes for personal (violent and sexual) offences than for property- (theft, burglary, criminal damage) or drug-related offences (e.g. Redondo et al. 2002). Fortunately, several meta-analytic reviews among the appended list are of particular note with reference to obtaining a fuller picture of the impact of interventions in reducing violent recidivism.

Dowden & Andrews (2000) integrated a series of 34 evaluations of interventions to reduce violence, yielding 52 effect-size tests. The target offence behaviours included general violence, and sexual and domestic assaults. Most (70%) of the studies included in this review focused primarily on work with adults. The overall mean effect size ($r$) was relatively low at $0.07$, though there was enormous heterogeneity in the findings: effect sizes ranged from a low of $-0.22$ to a high of $0.63$. The effect size for ‘human service interventions’, based on combining the principles defined by Gendreau & Andrews (1990) and now reconfigured as the risk–needs–responsivity model (Andrews et al. 2006) was $0.12$. Using the binomial effect-size display devised by Rosenthal & Rubin (1982), this figure corresponds to recidivism rates of 44 and 56% for experimental and control groups, respectively. Another important finding to emerge from this review was the evidence of a close correspondence between the number of criminogenic needs targeted in interventions and the associated effect size: a correlation coefficient of $+0.69$ ($p<0.001$).

More specific results have emerged from review of outcome studies with younger offenders, including those who have committed serious violent or sexual offences, as discovered in a more detailed meta-analysis by Lipsey & Wilson (1998). These authors integrated findings from a total of 200 studies, 83 of interventions delivered in residential settings and 117 delivered in the community. Lipsey and Wilson grouped types of interventions in broad categories defined by a combination of mean effect size and the consistency with which it was obtained. Intervention programmes in the most consistently effective category were found to have an average impact in reducing recidivism by 40% in community settings and 30% in custodial settings (Lipsey & Wilson 1998).

For community-based interventions, Lipsey & Wilson (1998) found that the largest mean effect sizes ($d$) were for structured individual counselling ($0.46$), interpersonal skills training ($0.44$) and behavioural programmes ($0.42$). For institutional-based methods, the largest mean effect sizes were for interpersonal skills training ($0.39$) and teaching family homes ($0.34$). Positive effects were obtained, but with less consistency, for the provision of ‘multiple services’ (sometimes called ‘service brokerage’, $d = +0.29$) in the community, and for behavioural programmes ($0.33$), community residential facilities ($0.28$) and multiple services ($0.20$) in institutional settings. Other types of intervention were either weaker or less consistent in their effects, or both. For a few interventions, notably deterrence-based initiatives, both this and the preceding review found effect sizes close to zero or negative.

There are three other meta-analyses of interventions to reduce aggressive or violent behaviour amongst youth. Wilson et al. (2003a) reviewed findings from 221 studies of interventions designed to reduce aggression in schools. The selected studies were carried out with participant samples ranging from pre-school to 13th grade (age 17–18 years), resulting in an aggregate sample of almost 56 000. Of the 522 comparisons possible between experimental and control groups, 34% were derived from randomized designs. The methods employed included competence training, with and without cognitive–behavioural components, classroom management techniques, counseling, separate streaming within schools, peer mediation, academic interventions, and varied combinations of the foregoing. Among the randomized designs, there was an overall effect size difference between experimental and control samples of 0.31 in favour of the former; the corresponding figure for non-randomized designs was 0.16. There were larger mean effect sizes for groups at opposite ends of the age distribution (less than or equal to 5 and more than or equal to 14 years) than for those in the middle age ranges. Both social competence training (with or without cognitive–behavioural components) and counselling yielded effect sizes in the $0.24–0.36$ range.

McCart et al. (2006) compared the relative effectiveness of behavioural parent training (BPT) and cognitive–behavioural therapies (CBT) in reducing aggression and other AB among young people under the age of 18. They found 41 studies of the former and 30 of the latter. The dependent variables were physical or verbal aggression, or delinquency. There was a mean effect size across all studies of $+0.40$ at post-test and $+0.22$ at follow-up (though the latter was based on only 17 studies). A direct comparison between the two approaches proved difficult as the mean age of those provided BPT was much lower than that of those provided CBT (5.44 versus 11.28 years). The comparison was thus restricted to 7 studies using the former and 21 using the latter that focused on the 6- to 12-year age group: respective mean effect sizes were $+0.45$ and $+0.23$. Across the full set of studies, the weighted post-treatment means were $+0.47$ and $+0.35$ for BPT and CBT, respectively, and in the 13 CBT studies for which follow-up data could be analysed there was a mean effect size of $+0.31$.

Garrido & Morales (2007) updated aspects of the Lipsey & Wilson (1998) review, though with a narrower focus on interventions provided in secure institutions only and confining the analysis to studies of groups defined as violent and chronic delinquents. Outcome
measures included both general ($k=15$) and serious ($k=30$) recidivism, the latter defined as comprising offences that led to reincarceration. There was a cumulative sample size 6658 and a median follow-up period of 18 months. The odds ratios for general and serious recidivism were 1.235 and 1.354, respectively, in favour of experimental/treatment groups. The latter effect size showed a surprising homogeneity. The authors sought to take account of sample attrition by conducting an intent-to-treat analysis, yielding an odds ratio of 1.307. All effect sizes are statistically significant.

For domestic violence, primarily consisting of assaults by males on female partners (designated intimate partner violence: Polaschek 2006), a meta-analysis has been reported by Babcock et al. (2004). These authors examined findings from 22 studies yielding (after elimination of outliers) 36 effect size tests; 17 of the studies were quasi-experiments and the remaining five ‘true’ experimental designs. For the Duluth model, using police reports as the outcome variable, the mean effect size ($d$) was +0.32 for quasi-experimental and +0.12 for experimental designs. For cognitive–behavioural methods, the mean effect size was +0.12 for quasi-experiments (no effect size could be computed for experimental designs). For partner reports, Duluth interventions had a mean effect size of +0.35 and CBT +0.29. Nevertheless Babcock et al. (2004, p. 1044) concluded that ‘...there is great room for improvement in our batterers’ treatment interventions’. It should be borne in mind that this is a singularly sensitive area in which to provide interventions and high attrition rates are common.

There have been six meta-analyses conducted to date of the effectiveness of interventions to reduce sexual recidivism (all listed in table 2). The most recent and comprehensive review (Lösel & Schmucker 2005) synthesized findings from 69 studies, covering a cumulative sample of 22 181 participants and including both medical and psychosocial treatments. From these findings Lösel and Schmucker were able to compute a total of 80 effect size tests. A majority (60%) of the studies consisted of non-equivalent group designs; for a further 19 equivalence was assumed, seven used statistical controls and six involved random allocation. Mean effect sizes across interventions, expressed as odds ratios (OR), were +1.70 for reductions in sexual recidivism, equivalent to a 37% reduction relative to comparison samples, +1.90 for violent recidivism (44% reduction) and +1.67 for general recidivism (31% reduction). The largest effects were for physical treatments (surgical castration, eight studies, $OR = 15.34$; hormonal medication, six studies, $OR = 3.08$). Some psychosocial interventions achieved significant effects (behavioural, 7 studies, $OR = 2.19$; cognitive–behavioural, 35 studies, $OR = 1.45$), while others (insight-oriented and therapeutic community approaches) had odds ratios not significantly different from 1. The mean effect size for cognitive–behavioural methods is lower than the odds ratio of 1.67 found in another review of sex offender treatment that focused solely on psychologically based interventions (Hanson et al. 2002).

Given the close association between violence and traits of callousness, low empathy, impulsivity and irresponsibility, a further set of meta-analytic reviews pertain to the effects of interventions with individuals classified as psychopathic. Unfortunately, the findings of the two available reviews in this area have not resolved the long-debated issue of treatability with this population. Salekin (2002) reviewed a series of 42 outcome studies; however, only 8 involved group comparison designs, and many others were single case reports, so while the latter may be clinically instructive any firmer conclusions must remain tentative at present. Of those studies that could be regarded as more robust, there were five studies of cognitive–behavioural therapy (CBT) incorporating a cumulative sample of 246 individuals. There were high effect sizes on intermediate outcome variables for several therapeutic approaches, including CBT, personal construct therapy and other approaches which ‘...addressed patients’ thoughts about themselves, others and society. Thus, they tended to directly treat some psychopathic traits’ (Salekin 2002, p. 93). Salekin also observed that there was a strong association between effect size and duration and intensity of treatment: interventions lasting less than six months were less likely to produce benefits than longer ones: where attendance was maintained for more than a year or delivered at a rate of more than four sessions per week, a considerably higher fraction of the samples benefitted. Working with civil psychiatric patients, Skeem et al. (2002) reported that those assessed as either psychopathic or ‘potentially psychopathic’ who received higher dosages of treatment (attended seven or more sessions) were approximately three times less likely to be involved in subsequent violence than those who received little or no treatment. This difference remained after controlling both for a number of background and clinical variables, and for treatment assignment.

The studies reviewed by Salekin (2002) predominantly reported increased levels of engagement in treatment, rather than treatment outcomes. More recently Tanasichuk & Wormith (2007) have reported a meta-analysis of outcome trials in this area. They located an initial total of 21 studies yielding 50 effect size estimates (cumulative sample $n=5550$). In comparisons between those designated as psychopaths and samples of non-psychopaths, the former consistently showed higher general, violent and sexual recidivism, more AB, higher levels of substance abuse and spent significantly less time in treatment. In three studies where comparisons were possible between treated and untreated psychopaths, there were no significant differences in general or violent recidivism; other types of comparisons were not feasible given the available data. However, contrary to the findings of some earlier research (Rice et al. 1992), there was no evidence that treatment made psychopaths worse, and a study by Wong et al. (2006) found that psychopaths completing Aggressive Behavioural Control, a violence risk reduction programme at the Saskatoon Regional Psychiatric Centre, subsequently committed less serious offences than matched controls. While there was no significant difference in total reconvictions, ‘the mean length of the sum of the aggregate sentences was halved and the longest aggregated sentence was more than halved’ (Wong et al. 2006, p. 3).
Outcome studies in this area are relatively few in number and suffer from poor designs, leading Tanasichuk & Wormith (2007) to recommend that more matched control group studies should be conducted. Other authors researching this area have remained unconvinced that evidence concerning therapeutic change among diagnosed psychopathies is likely to be forthcoming (Harris & Rice 2006). Against such ‘therapeutic nihilism’ as it has been called, it is equally plausible that appropriate interventions adopting ‘best practice’ formulae have not yet been devised for working with this group. Long-term monitoring of the diagnostic status of those classified as suffering from personality disorders, including psychopathy, indicates that reductions in symptom severity occur. Serin (1995) noted longitudinal evidence concerning decreasing portions of study samples retaining diagnostic features. This is amplified by Sanislow & McGlashan (1998) who reviewed 44 studies of the ‘natural course’ of personality disorders (PD) including antisocial PD. Rather than finding a fixed, immutable pattern as was previously expected, this review showed a pattern of changeability over time.

If individuals are assessed as having more extreme scores on measures of various criminogenic risk factors, it is not surprising to find that more complex, intensive, multi-modal methods are required to engender attitudinal or behavioural change. Furthermore, two studies have shown that it is possible to improve the ability of children and adolescents with callous–unemotional traits (considered a precursor component of the emergence of psychopathy) to recognize facial expressions of fear, by redirecting their attention to other people’s eyes (Richell et al. 2003; Dadds et al. 2006). If this is feasible for children thought to be at long-term risk of developing psychopathy, it suggests that the variables associated with aggression can be ameliorated. It also indicates a possible method of doing so that could be modified and incorporated in multi-modal treatments.

Finally, Leitner et al. (2006) reported a systematic review of risk assessment and violence prevention in the field of forensic mental health. Of 228 182 citations initially retrieved, 299 evaluations that employed some form of between-groups statistical analyses were retained after screening. Among these studies, 54.5% were of pharmacological agents, 29.8% of psychosocial interventions, 6.7% some combination of the former two and 9% were classed as ‘other interventions’ (e.g. organizational changes in ward management procedures). Within these general categories, the proportions of studies with unequivocally positive outcomes were 80.4, 79.8, 85 and 66.7%, respectively. Hence in general, there were fairly high success rates of interventions of a variety of types for reducing violence in this client group.

Currently, Whittington et al (in preparation) are undertaking meta-analysis of the best controlled trials in this dataset, to be reported in a future paper. Of 112 studies of psychosocial interventions that were located, only 28 employed randomized designs. This differs somewhat from the results of a scoping survey (Cure et al. 2005) which suggested that over 700 interventions to reduce aggression have been tested in 300 controlled trials, including 328 ‘named talking therapies’. The target populations in the studies are very varied and include diagnoses of schizophrenia, attention-deficit–hyperactivity disorder, individuals who have committed domestic violence or child abuse, prisoners and social care residents or offenders with learning disabilities. Similarly, there is a wide variety of intervention methods, including socialization games, social skills training, relaxation training, anger management, behaviour modification, cognitive–behavioural therapy, motivational interviewing, parental skills training, family therapy, group psychotherapy and psycho-dynamic therapy. Unfortunately, most methods are represented by a single study only, and studies also employ a miscellany of outcome measures, including self-report and psychometric scales, observational measures, and official data sources such as criminal reconvictions. Initial effect size analyses (using Cohen’s $d$) show variations from +0.01 to +1.40.

Overall, on the basis of the series of 11 meta-analyses just discussed, addressing violent behaviour in general or specific forms of it, there are numerous positive outcomes. These permit reasonable confidence in the broad conclusion that it is possible to reduce violent recidivism by systematic and carefully designed intervention. As with offender treatment considered more broadly and apart from the findings obtained for physical treatment of sexual offending, the most consistent outcome effects are for a collection of methods derived from the cognitive social learning model (behavioural, cognitive, interpersonal and problem-solving training methods). However, a considerable need remains for more, better controlled outcome studies to test more refined hypotheses concerning the relationships between intervention methods, offence typologies, participant characteristics, delivery settings and other variables, and to address the enduring problem of transfer of findings to routine practice in criminal justice or mental health services.

6. ILLUSTRATIVE APPLICATIONS AND OUTCOMES

Meta-analyses of treatment studies are of course to some extent abstractions. Subsuming information from various sources, they are removed from direct clinical work or service delivery that is being evaluated in each of the studies. In order to portray more accurately the nature of interventions tested in these reviews, this section provides some illustrative outlines of specific programmes, methods and corresponding results.

(a) Reactive (emotional) aggression: anger management

Numerous behavioural and cognitive methods have been applied to angry aggression, perhaps most notably the model of anger reactions and their management developed by Novaco (1975, 1997, 2007), which has had a seminal effect in many clinical locations, including work with offenders. The model describes the interdependence of cognitive appraisal, emotional arousal and angry response in a manner that enables individuals to both understand their experiences of
anger, and acquire a framework and techniques for regaining control where they may have lost it.

There are several meta-analytic reviews of anger management/anger control, covering work with a wide range of populations. In an early review, Edmondson & Conger (1996) carried out a synthesis of 18 studies. Most had small samples and the total contact time (session length multiplied by number of sessions) was fairly limited, usually amounting to only 6–8 hours. The effect sizes were remarkably high, though they varied according to the outcome target. The authors suggested that the choice of anger treatment should depend on the specific types of anger problems experienced by the participants.

More recently DiGiuseppe & Tafrate (2003) reported a meta-analysis of 50 studies (with a total sample of 1841 adult participants and 81 effect-size tests). Many different types of therapy were evaluated in the studies reviewed: including self-instructional training; cognitive restructuring; problem-solving; relaxation training; systematic desensitization; exposure therapy; behavioural self-management; anxiety management training; and various combinations of the foregoing. Their review found positive effects of interventions on both expression of anger and aggressive behaviour, and the maintenance of gains over time. The mean effect size for anger control was $+0.71$ and for aggression-related outcomes $+1.16$, a large effect.

A separate review has also been reported, focusing on anger-based interventions for children and adolescents ranging in age from 6 to 18 years (Sukhodolsky et al. 2004). There were 40 studies included, 80% of which involved random assignment, generating 173 outcome tests and a cumulative sample of 1953. The mean effect size ($d$) for anger experience was $+0.47$ and that for physical aggression was $+0.63$.

Additional reviews have followed. Gansle (2005) synthesized findings from 20 studies conducted in schools; 75% used random assignment. The mean effect size across studies for externalizing anger problems was $+0.54$ at post-test and $+0.53$ at follow-up. Another review was restricted to randomized trials only (Del Vecchio & O'Leary 2004). There were 23 studies with a total sample of 1730 participants. Mean effect sizes for four categories of intervention represented were: for relaxation training alone, $+0.90$; for cognitive therapy alone, $+0.82$; for cognitive–behavioural therapy, $+0.68$; and for a mixed group of other treatments, $+0.61$. The effect sizes were such that ‘...as per this analysis, hundreds of additional studies averaging negligible results would be needed to reduce these findings to negligible levels, alleviating the effect of the file drawer problem’ (Del Vecchio & O'Leary 2004, pp. 25–26).

Overall, the results for interventions designed to reduce anger or render it more manageable, and decrease resultant aggression, appear impressive. Effect sizes are in the moderate-to-large range and there is a healthy representation of well-controlled trials. However, with some notable exceptions, the majority of participants in the studies reviewed are general population samples reporting anger as a problem and seeking help to manage it. In work with violent offenders, while there have been some valuable outcomes, overall results are less consistent. It should be emphasized that while some violent offences may be directly caused by the loss of control over anger, it is difficult to ascertain how much this contributes to violent behaviour in general (Polaschek & Reynolds 2004). It appears essential to assess a pattern of offending (using functional analysis and case formulation) prior to allocating individuals to programmes. Polaschek (2006) has noted the inappropriateness of anger management alone for individuals exhibiting either over-controlled or instrumental aggression, though it could constitute an element in a multi-modal programme focused on such problems.

In penal settings, one of the most extensive applications of anger control programmes was carried out in Canadian prisons. The programme consisted of 25 two-hour sessions, offered between two and five times per week, for groups varying between 4 and 10 prisoners. Dowden et al. (1999) reported a 3-year follow-up of 110 programme participants and matched controls. For lower-risk cases, there was no impact on levels of reoffending. For high-risk cases, however, there was a 69% reduction in general recidivism, and an 86% reduction in violent reoffending, one of the largest effects reported in the literature. Dowden & Serin (2001) reported a more searching analysis of the follow-up data taking account of performance measures (indicators of extent of participation in the programme). There was a correlation of $+0.32$ between these measures and recidivism outcomes, making them the strongest predictor of recidivism. At the follow-up point, there were significant differences ($p < 0.001$) between treatment completers, untreated comparisons and programme dropouts, with respective rates of general recidivism 10, 30 and 52%, and of violent recidivism 5, 17 and 40%. These findings suggest that there may be an underlying treatment effect, but this cannot be ascertained with confidence given the higher rate of recidivism among the dropouts than among the controls.

Implementation of anger management training in penal settings has not proved to be uniformly successful. In some instances treatment gains have been marginal. Howells et al. (2002) carried out a large-scale evaluation of anger management programmes in several prisons in Australia. On the basis of their study, the authors made a number of observations: while arguing that anger management interventions should be continued, they recommended moving away from ‘blanket delivery’ of programmes. The same authors have also drawn attention to the importance of readiness for change in assigning prisoners to anger control sessions (Howells & Day 2003).

(b) Interventions with young offenders

In the extensive meta-analytic review reported by Lipton & Wilson (1998) described above, several intervention methods emerged as most consistently yielding positive outcomes with regard to reduced rates of violent and sexual recidivism among young offenders.

*Phil. Trans. R. Soc. B* (2008)
(i) **Interpersonal skills training**

This comprises a series of exercises designed to improve participants’ skills in interacting with others. Working in a small group, individuals identify situations in which they are uncertain how to act or which they sometimes mishandle. Suitable ways of behaving in the situation are discussed, practised using role play and shaped via behavioural rehearsal and feedback. An example of this is a study by Chandler (1973) on training in ‘perspective taking’ skills.

(ii) **Structured individual counselling**

The most widely used format of counselling as a relatively unstructured, person-centred, non-directive activity has not emerged as an effective means of reducing offender recidivism. However, more structured approaches based, for example, on reality therapy, problem-solving or multi-modal frameworks of rewards for progress towards modifying visors compose a list of problem behaviours and a contracts, where individual offenders and their super-

In work with offenders, this has included contingency this intervention has generally demonstrated positive outcomes with very challenging young people.

(iii) **Behavioural interventions**

In work with offenders, this included contingency contracts, where individual offenders and their supervisors compose a list of problem behaviours and a system of rewards for progress towards modifying them, in conjunction with behavioural training procedures such as modelling and graduated practice, and cognitive and problem-solving skills training. For example, using behavioural methods in work with families, large reductions in juvenile offending were obtained over follow-ups of $2\frac{1}{2}$–3½ years (Klein et al. 1977; and see Gordon 2002).

(iv) **Teaching family homes**

These are residential units or group homes in which especially trained adults work in pairs as ‘teaching parents’. Their role is to develop positive working alliances with residents, impart a range of interpersonal and self-management skills and provide counselling and advocacy services. Young people can continue to attend school and return to their homes of origin at weekends (see Kirigin et al. 1982).

(c) **Cognitive skills programmes for adults**

Cognitive skills programmes are so called because their objectives and the methods they employ are directed towards helping the participants to acquire new capacities for thinking about and solving their problems, particularly in the interpersonal domain. They draw on earlier work such as that of Chandler (1973) or Platt et al. (1980) (for background, see McGuire 2005).

Programmes of this type are derived from the cognitive model of offender rehabilitation (Ross & Fabiano 1985), a variant of social learning theory placing a particular accent on cognitive skills. These denote the capacity when faced with a personal difficulty to engage in a sequence of activities including problem identification, generating alternative solutions, means-end thinking, anticipating consequences of actions and perspective taking. Ordinarily, such skill acquisition occurs naturally during development, but that process is a function of appropriate learning opportunities. Conversely, the absence of such skills is held to constitute a risk factor for AB including resort to aggression and violence.

The most widely disseminated programme derived from the model, entitled Reasoning and Rehabilitation (R&R), consists of a series of 38 two-hour sessions delivered on a group basis by special trained tutors (Antonowicz 2005). Its constituent materials are organized into a sequence of interlinked modules focusing on problem-solving, social interaction, impulse control and self-management, negotiation and conflict resolution, and critical thinking.

In a large-scale evaluation for Correctional Services Canada, with a sizeable sample of federally sentenced prisoners ($n=1444$), there was a reduction in recidivism of 36.4% among those completing the programme when compared with controls (Robinson 1995; Robinson & Porporino 2001). Of specific relevance here, these effects were moderated by offence type: prisoners with records of violent, sexual and substance-related offending were less likely to be reconvicted than those with histories of property crimes. Both R&R and other cognitive skills programmes have emerged as beneficial from focused meta-analytic reviews (Wilson et al. 2005a; Tong & Farrington 2006).

However, system-wide implementation of R&R in prisons in England and Wales yielded mixed and predominantly negative results (Friendship et al. 2002; Cann et al. 2003; Falshaw et al. 2003).

Results for probation-based versions of cognitive skills interventions have been generally more positive. There are significant reductions in actual 2-year recidivism rates below expected rates across most types of programme (Hollis 2007). Where it has been possible to enter prior levels of risk of reconviction among programme participants, non-completers and comparison samples, as variables in logistic regression analyses, offenders completing programmes have been found to have significantly lower rates of subsequent recidivism (Palmer et al. 2007). In order to take account of between-group differences in the likelihood of programme non-completion, such data can also be analysed using propensity scores, separating variables likely to influence completion from the potential impact of programmes on recidivism as the dependent variable. Where this has been done, again there is evidence of a link between programme completion and a significant reduction in criminal recidivism (McGuire et al. 2008). Within this, reductions are observed in rates of violence alongside other types of offending, though the relationship between this and programme completion has not been analysed separately.

(d) **Cognitive self-change**

Cognitive skills programmes are designed to impart to their participants a series of cognitive, interpersonal and self-management skills, limitations or deficits in which are thought to have contributed to the occurrence of acts of crime. An alternative approach is to
consider such acts as arising from cognitive distortions held by the offender: beliefs or assumptions that are directly conducive to antisocial acts.

A programme of this type was developed within an adult prison establishment in the Vermont Department of Corrections, USA. Bush (1995) had described the rationale for the programme and its mode of delivery. Sessions were run within a separate unit inside the prison; groups of between 5 and 10 prisoners with histories of violent offending met three to five times per week. In each session, one prisoner was asked to describe an incident in which he had been involved and to furnish a thinking report. This is a detailed record of thoughts and feelings before, during and after a violent act, in which participants seek to obtain a fuller understanding of the factors that have influenced them. Groups collaborated in identifying criminogenic thought patterns; then generating new thoughts or practising skills that would make violent behaviour less likely, a procedure entitled self-risk management.

Henning & Frueh (1996) reported a 2-year follow-up of 55 prisoners who attended this programme for an average of 9.8 months, compared with an appropriately matched sample of 141 non-attenders. There was a significant difference in the respective recidivism rates of the two groups (50 versus 71%). The follow-up analysis also showed that members of the experimental group survived significantly longer in the community before committing new offences.

(e) Other multi-modal interventions for violence
A more elaborate multi-modal programme to address violent offending is described by Cortoni et al. (2006). The Violence Prevention Programme (VPP) consists of many different elements and was based on a model that incorporated motivational enhancement, behavioural change methods, a focus on aggressive beliefs, cognitive distortions, arousal management, impulsivity, conflict resolution, problem-solving, empathy enhancement and relapse prevention (Serin & Preston 2001).

The evaluation study consisted of two samples: 500 VPP participants and 466 prisoners in a matched comparison group. Members of both groups were in custody during the period 1999–2004; 66.6% of the former group completed the programme. The comparison sample was matched to the experimental group using propensity scores defined with reference to the probability of receiving treatment. Evaluation data included numbers of institutional incidents (disciplinary infractions) as a proximate outcome, with general and violent recidivism as follow-up outcomes after discharge into the community, taking account of time at risk. Those who completed the programme had a uniformly lower failure rate (on all measures) than the comparison sample, notably a violent recidivism rate of 8.5% when compared with that for the non-completers of 24.5% and for the comparison group of 21.8%. Expressed as risk ratios using Cox regression, the comparison group had a 1.36 times greater risk of any failure, and a 2.10 times higher rate of violent recidivism, than completers. For non-completers, the rate of violent recidivism was 4.25 times that of completers. Similar findings were obtained for Aboriginal offenders when data for that group were analysed separately.

Possible confounding variables included a difference between the groups in terms of motivation for intervention which was higher among the experimental than the comparison sample; and the fact that a slightly higher proportion of treatment-group members had completed other violence-related programmes prior to embarking on the VPP.

Other multi-modal cognitive skills group programmes that have demonstrated success in reducing violence include the Montgomery House Violence Prevention Project and the Violence Prevention Unit (VPU) in New Zealand (Polaschek & Reynolds 2004; Polaschek et al. 2005; Polaschek 2006). While the study of the former was statistically underpowered, there was a large observed reduction in violent reoffending for treated offenders and minimal change in a matched untreated control group. In a 5-year follow-up evaluation, of 64 men in the treated group, 33 committed new offences and of the same number in the control group, 51 committed new offences. When the outcomes were adjusted for the rate of pre-programme violence, the recidivism rates were 25 and 42% for the experimental and control samples, respectively (Montgomery House 2007). For the prison-based VPU, the first 22 completers were compared with a matched untreated sample drawn from a national database. At a minimum 2-year follow-up, 32% of the completers had a violent reconviction alongside 63% for the comparison group; survival analysis found an effect-size difference between the two groups of +0.41 in ‘days to failure’.

Finally, Aggression Replacement Training (ART) is a multi-modal programme that employs methods of social skills, anger management and moral reasoning training in an integrated, 30-session format (Goldstein & Glick 2001). Evaluations of its use with young offenders, while based on small samples and in non-equivalent designs, have found positive effects, but are reported in terms of general rearrest rates rather than violent recidivism. Aos et al. (2001) summarized four studies with adjusted effect sizes ranging from 0.07 to 0.26. On the basis of unpublished Home Office data concerning the use of ART with adults on probation, McGuire & Clark (2004) reported that programme completers (n=113) had a significantly lower rate of recidivism at a 12-month follow-up than matched comparison samples allocated to other types of probation supervision. While the predicted and actual reconviction rates for the comparison sample were, respectively, 37.9 and 34.5%, the corresponding figures for the ART participants were 34.6 and 20.4%. Predicted and actual rates of those allocated to the programme, or to other probation interventions, whose probation orders were revoked, were uniformly higher. Thus the ART completer group was the only one to show a significant reduction in reconviction below predicted levels. Again, however, the use of a quasi-experimental design leaves the results open to more than one interpretation.

Using an ‘intent-to-treat’ analysis, with a comparison sample formed on the basis of one-to-one matching, Hatcher et al. (in press) found a reduction
of 13.3% in reconviction among those allocated to ART relative to controls. Compared with their matched counterparts, programme completers had a reconviction rate 15.1% lower, and completion of the programme was associated with a 78% improvement in survival times. This was, however, a fairly short follow-up (10 months) and the completion rate for the programme during the study period was rather low; these are nevertheless reasonably promising results.

7. CONCEPTUALIZING VIOLENT BEHAVIOUR

Given the multiplicity of factors shown to influence violent conduct and the variety of interventions with some supportive evidence of effectiveness, the construction of a viable model of causation and intervention remains an elusive goal. Arguably, a great deal is known about separate risk factors and permutations of them that are associated with delinquency and crime in general (Farrington 2007) or with a repetitive pattern of violence in particular (e.g. Herrenkohl et al. 2000; Surgeon General 2001; Farrington 2003). Not unreasonably, some authors (Rutter 2003; Moffit & Caspi 2006) have expressed dissatisfaction with the risk factors approach and the absence of coherent accounts of causal pathways confirmed by empirical findings. However, models of functional relationships between factors operating at different levels (biological, psychological, social) and different developmental stages (infancy, middle childhood, adolescence) have been proposed (e.g. Dodge & Pettit 2003; LeBlanc 2006) and are capable of accounting at least in general terms for a sizeable proportion of the data currently available.

There is a consensus that violence is only likely to be comprehensible within probabilistic rather than deterministic causal models, an assumption pivotal to the risk factors approach. But the extent to which this can be developed into a formal theory remains questionable. An ineluctable problem arising in all research in this field is that most discrete risk factors or other variables account for only a small proportion of the measured variance in violent behaviours, underlining the necessity of a multifactorial model. What should be the nature of such a model?

Tackling the task of theory construction, some formerly intractable problems are addressed and some fundamental principles are enunciated by Dodge & Pettit (2003). One is that of equifinality, whereby ‘...the same antisocial outcome can accrue from disparate sources’ (p. 354). Another is that of multifinality, whereby ‘...specific risk factors can be associated with a variety of outcomes’ (p. 354). This is to some extent reminiscent of the problem known in the philosophy of science as ‘underdetermination of theory’, wherein the available data are potentially compatible with more than one of the theoretical formulations on offer (Klee 1997).

But it may be that, given the nature of the problem, data that would allow a critical choice between theories, or even allow differential apportioning of causal weight to diverse classes of variable, will not be obtainable. Any theoretical model that is designed to account for the available data at anything other than a macroscopic level may be unlikely to succeed. A possible route out of this dilemma is offered in the philosophical work of Mackie (1980) concerning causality. In addition to the familiar Humean categories of necessary and sufficient causes of events, Mackie specifies a third category of factors entitled inus conditions. This acronym refers to an insufficient, non-redundant part of an unnecessary but sufficient condition. The possible meaning of this is represented diagrammatically in figure 1.

Thus there are different sets of factors (denoted A, ... D, ... in figure 1) that influence the likelihood of a violent act (or pattern of perpetrating such acts; denoted P). They are customarily grouped under familiar headings (e.g. neurobiological, cognitive, social) and within each there is a potentially extensive list of defined independent variables (e.g. monoamine oxidase A polymorphism, suppressed amygdala activity, dysexecutive syndrome, callousness, impulsivity, attributions of hostility, coercive family process, peer pressure, social inequality, media images, provocations, intoxication). Given any instance or sample of instances of persistent aggressiveness, there could be different patterns of influence at work yielding the same outcome, operating along multiple pathways (Dahlberg & Potter 2001). Each of the aforementioned factors alone (or any of numerous others) may be an insufficient cause of aggression by itself; but it cannot be dispensed with (it is non-redundant) as part of a larger set of causes, each of which is unnecessary to the observed outcome (i.e. it could occur in their absence) but may nevertheless collectively be sufficient to cause it in some cases. Where the independent variables in a research study account for only a small proportion of the variance in the dependent variable (the standard scenario), or where the findings of different studies contradict each other, it may be that this is due to them each having measured some inus conditions and not others.

Returning to the evidence reviewed in this paper concerning effective interventions, a core question that remains is how those changes that engender reductions in aggression and violence occur (‘treatment theory’; Polaschek 2006). While numerous variables operate during such a process, arguably what they share in common is an increase in the individual’s capacity for, and motivation to engage in, self-regulation of cognitive

Figure 1. inus conditions and probabilistic models of violence causation.

Phil. Trans. R. Soc. B (2008)
and emotional processes and of the interaction between them. Longitudinal research by Bandura et al. (2001, 2003; Caprara et al. 2002) using structural equation models has revealed significant linkages between low perceived self-regulatory efficacy and patterns of violent conduct in young people followed up between the ages of 11 and 19. Working with a different age group in very different circumstances (recidivist adult offenders who had already served several prison sentences and were reincarcerated again following a new offence), Zamble & Quinsey (1997) found that poor self-regulation, misdirected attempts at problem-solving and failure to manage negative emotions were connected with the occurrence of reoffences.

There are invaluable integrative frameworks within which these observations might be subsumed in striving towards a properly articulated theory. They include the biopsychosocial model proposed by Dodge & Pettit (2003), deploying concepts of nonlinear transactional relationships between key variables. Another example is the work of LeBlanc (2006) who draws extensively on control theory in criminology in constructing a developmental model of interactions between personal and social controls in the genesis of deviant behaviour. A still broader conceptual framework (applicable to clinical problems beyond that of violence) is the ‘hot/cool system’ analysis of Metcalfe & Mischel (1999).

Arguably within each of these perspectives, however, ultimate behavioural outcomes must be mediated by psychological variables (Kinderman 2005). Whatever the distal factors may be that lead through multiple developmental pathways to aggressive propensity, in the overwhelming majority of cases these interact with current circumstances and are expressed through experiential (including cognitive and emotional) processes and events. Kinderman (2005) has forwarded a general conceptual scheme depicting these relationships, shown in a modified form in figure 2.

According to Dodge & Pettit (2003), ‘life experiences with parents, peers and social institutions mediate, at least partially, the effects of biological predispositions and sociocultural contexts’ (2003, p. 357) in ways that lead to disordered conduct. They also assign a core role to ‘agentic emotional and cognitive processes’ that are ‘postulated as the crucial factors that mediate the relation between risk factors and conduct problems’ (2003, p. 361). Kinderman (2005) argued that by ascribing equivalent status to biological, psychological and social factors in causation, the biopsychosocial model pays insufficient attention to the role of psychological processes as a ‘final common pathway’ in mental disorder. Such a view is equally applicable to understanding violent behaviour, in that regardless of the distal origins of contributory factors, to result in overt actions they must be expressed through psychological processes. Similarly, the basis of effective violence reduction should be grounded in addressing aspects of those processes in intervention efforts.

8. CONCLUSIONS

On the basis of the present review, the following conclusions are offered regarding the prospects of being able to reduce established patterns of aggression and personal violence. First, there are large amounts of evidence showing that it is possible to reduce the rate of occurrence of these problems among individuals who have been identified as manifesting them. There are methodological weaknesses within segments of that evidence and inconsistencies among reported outcomes, but there are sufficient indications to detect several encouraging trends. Emotional self-management, interpersonal skills, social problem-solving and allied training approaches show mainly positive effects with a reasonably high degree of reliability. Findings are weaker with respect to domestic violence and less consistent with reference to prison-based programmes. Given our present state of knowledge, findings to date of treatment non-responsiveness among some groups are arguably best interpreted as ‘absence of evidence rather than evidence of absence’.

Second, therefore, with regard to almost all issues, there is a need for more and better quality research if this is not to remain an ‘underdeveloped corner of offender rehabilitation’ (Polaschek 2006, p. 145). It is important to maintain pluralism of research designs to test both internal and external validities. That is, there need to be more randomized experiments to test specific hypotheses. But it is important not to abandon practical trials that will attest to the usefulness of methods in routine service delivery: quasi-experimental studies can yield valuable information that may be more easily transferred to practical settings, provided researchers adhere to the TREND guidelines (Des Jarlais et al. 2004; and see Hollin 2008). Given the complexity of the problem, it appears advisable to research multi-modal interventions only, but the contribution of separate components could be evaluated in dismantling designs.

Third, in relation to practice, for most programmatic interventions in extant use, it is almost certainly necessary to increase the duration and intensity of treatment (‘dosage’) above presently inadequate levels if intended effects are to be obtained. In criminal justice services, there is a need to improve targeting, preferably applying the risk-needs–responsivity framework (Andrews et al. 2006) as currently the best validated model. In implementing this at an individual case management level, it is imperative to employ functional analysis/case formulation for treatment allocation decisions.

Phil. Trans. R. Soc. B (2008)


