

Implicit Measurement of Violence-Related Cognitions

Stephen Blumenthal, Nicola S. Gray, Richard Shuker, Heather Wood, Peter Fonagy, Miranda Allonby, Alison Flynn, Tanja Takala and Robert J. Snowden

Author Notes:

Stephen Blumenthal, Heather Wood, Miranda Allonby, Alison Flynn, Tanja Takala, Portman Clinic, London, UK

Nicola S. Gray, School of Psychology, Swansea University, UK and Caswell Clinic, Abertawe Bro Morgannwg University Health Board, UK.

Richard Sucker, HMP Grendon, [Grendon Underwood](#), [Buckinghamshire](#), UK.

Peter Fonagy, University College London. London, UK

Robert J. Snowden, School of Psychology, Cardiff University, UK

Correspondence concerning this article should be addressed to Stephen Blumenthal:

Portman Clinic, Tavistock & Portman NHS Foundation Trust, 8 Fitzjohn's Avenue, London NW3 5NA. sblumenthal@tavi-port.nhs.uk

Acknowledgements

Our thanks to the staff and inmates of HMP Grendon for their help in this study.

Abstract

Objective: Violence-related cognitions may underpin many acts of violence, but explicit self-report measures of these cognitions may be inadequate to assess them fully due to their unconscious nature or due to deliberate dissimulation. We designed three versions of the implicit association test (IAT) that separately examined violence-related associations to valence (good vs bad), hedonic-value (enjoy vs dislike) and arousal (exciting vs boring) and examined if these were associated with greater levels of past criminal activity and self-reported violence. **Method:** The three IATs were administered to an offender sample ($N = 108$) with a history of serious offending and a diagnosis of Personality Disorder. Violence was assessed via the Proactive and Reactive Aggression Questionnaire and by examination of previous convictions. **Results:** The IATs showed good to moderate reliability. The valence-IAT did not show any significant correlations to the measures of previous violence. The hedonic value-IAT showed positive relationships with official records of convictions, especially among participants without a conviction for homicide. The arousal-IAT was positively related to self-reported aggression in those without a conviction for homicide. **Conclusions:** The results show some promise that indirect techniques may be able to measure violence-related cognitions.

Keywords: implicit association test, reactive aggression, proactive aggression, personality disorder.

Implicit Measurement of Violence-Related Cognitions

In 2013 law enforcement made nearly 0.5 million arrests in the USA for violent crimes, and there were over 14,000 murders. In England and Wales there were an estimated 1.3 million incidents of violence, including 574 murders, in the period 2014/2015 (Office for National Statistics, 2016). Clearly, understanding the motivations for violent acts can serve as a starting point for intervention programmes that aim to reduce violence at either a societal level or at the individual level. To this end we aimed to test whether novel indirect measures of violence-related cognitions were related to levels of violence both via self-report and official records.

Violence-related Cognitions

Many studies have sought to measure cognitions related to violence that might underpin violent actions (Bowes & McMurrin, 2013). Polaschek, Calvert, and Gannon (2009) analysed transcripts of offenders describing their crimes and noted a series of common violence-related cognitions, such as 'violence is normal', 'beat or be beaten', and that violence is 'uncontrollable'. Such attitudes and thoughts about violence and aggression may be associated with risk of violent action and could play an important part in the aetiology and maintenance of violent behaviour. The ability to be able to illustrate changes in violence-related cognitions may be indicative of therapeutic progress in individuals with a history of violence, or, indeed, signal an escalating risk of violence in those offenders managed by probation services within the community.

Measurement of violence-related cognitions is, however, not straightforward. By far the most commonly used method is to directly ask the individual about such attitudes via interview or questionnaire (Buss & Perry, 1992). However, this method has problems. First,

many cognitive processes are simply not available to introspection and/or consciousness (Wilson & Dunn, 2004). It has been suggested that some violent offenders have ‘implicit’ theories of the world that underpin their behaviours (Polaschek et al., 2009; Weldon & Gilchrist, 2012), but that these ‘implicit’ theories are not available to conscious access or to explicit report. Furthermore, individuals may deliberately distort, minimise, or deny these cognitions. Hence, it is difficult to rely on self-report of such sensitive issues in populations that may have good reason to dissimulate these attitudes or motivations to violence.

Indirect Measurement of Cognitions

Faced with the problems related to direct, or explicit, measurement of negative, or anti-social, cognitions, researchers and clinicians have begun to use indirect or implicit methods (for a review see Snowden & Gray, 2010). One popular method is the Implicit Association Task (IAT: Greenwald, McGhee, & Schwartz, 1998). In an IAT the participant classifies two sets of stimuli according to a pre-set scheme. For example, Gray, MacCulloch, Smith, Morris, and Snowden (2003) had offenders classify words into categories of ‘violent’ or ‘not-violent’, and other words into ‘pleasant’ or ‘unpleasant’. This task was done in two blocks. In one block the ‘violent’ and ‘pleasant’ stimuli had to be classified onto the same response button, while in the other block of trials the ‘violence’ and ‘unpleasant’ stimuli shared the same response button. Most people were fast and accurate when the task was arranged so that ‘violent’ and ‘unpleasant’ went together but were slow when ‘violent’ and ‘pleasant’ went together. This was interpreted as showing that most people have an automatic cognitive association that violence is bad. Gray et al. (2003) showed that this effect was significantly weakened in those who had committed murder and were classified as psychopathic (see also Snowden, MacCulloch, Smith, Morris, & Gray, 2004). This result has been replicated in a Central-American sample (Ostrosky-Solis, Rebollar, Garcia, &

Villalpando, 2009), with male adolescent offenders with conduct disorder (Olivera-La Rosa et al., in press), and has been shown to be associated with greater levels of intimate partner violence and poor treatment outcome in offenders (Eckhardt, Samper, Suhr, & Holtzworth-Munroe, 2012). Other studies have shown that a strong negative implicit view of violence (as measured by the violence-IAT) is associated with increased pro-social behaviours and attitudes (Zwets et al., 2015) and with a greater history of trauma (Bluemke et al., 2017).

Indirect measures of cognitions have been shown to be predictive of behavior in many settings (see Perugini, Richetin, & Zogmaister, 2010) and have been suggested to be particularly predictive of behaviors which are spontaneous or where the person has little time or cognitive resources to weigh up decisions/actions (e.g., Friese, Wanke, & Plessner, 2006). Thus, violence-IATs may be particularly valuable in measuring attitudes to violence when the person is placed under pressure to act, or under threat.

Valence, Hedonic-value and Arousal.

Indirect measures of violence cognitions have so far examined associations between violence and valence (good vs bad). However, violence could be evaluated among many other dimensions. As an example, let us consider attitudes to cigarettes. Even most committed smokers would probably concede that cigarettes are bad (Chassin, Presson, Sherman, & Edwards, 1991). Hence, the 'classic' valence-IAT that compares the target concept (smoking) to a valence dimension (good vs bad) might be expected to reveal negative attitudes for both smokers and non-smokers (see Huijding, de Jong, Wiers, & Verkppeijen, 2005). However, these same smokers might also say that they enjoy or like cigarettes despite their ratings that cigarettes are bad. Hence, one can conceptually distinguish hedonic effects (whether one enjoys/likes something) from an evaluation of its valence (whether it is good or bad). Finally,

an ex-smoker might regard cigarettes as ‘bad’ and might even say that they no longer ‘enjoy’ cigarettes, although they may still want or crave cigarettes (want/crave).

We note that a distinction between ‘liking’ and ‘wanting’ has proved most valuable in the study of drug-addiction (incentive salience theory; see Robinson & Berridge, 2003) where there can be a clear distinction between changes in how much someone enjoys a drug, and how much they crave or want that drug. However, whilst there are clearly some (interesting) parallels between this theory of drug addiction and attitudes to violence, we do not tie our hypothesising to the incentive salience theory that has spawned much research on liking and wanting (Robinson & Berridge, 2003), i.e., we do not claim that violence has similar properties to addictive drugs.

Reasoning in a similar manner for violence as we did for cigarettes, it would seem that most people would volunteer that violence was a ‘bad’ thing – and both explicit and implicit measures support this, even in offender populations (Snowden et al., 2004). Despite this general view that violence is bad, many people pay substantial amounts of money to see violent sporting events, such as boxing matches, where the violence is central to the entertainment. It would seem likely that such individuals would report that they ‘enjoy or like’ such events. Finally, for some individuals that have perpetrated violence, this violence may evoke a very high level of arousal or excitement. Hence, it may be possible to examine the amount of arousal that a person associates with violence. We hypothesised that any person’s attitude to violence would contain components of valence (whether they believe violence to be a good/bad thing), hedonic-value (how much they enjoy/dislike violence), and arousal (how excited they get by violence). It also seems possible that these cognitions might well be important in the aetiology and/or maintenance of violent behaviour. We reasoned that such attitudes might be hard to elicit from self-report (explicit) measures as these are prone to

dissimulation, with both faking good and faking bad (e.g., enhancing one's macho presentation and image) occurring in an offender population. Hence, we argue that indirect measures are needed to measure them.

To date, there has been no attempt to look at motivations to violence using implicit techniques that contrast valence, enjoyment, and arousal (though see Tibboel, De Houwer, and Van Bockstaele (2015) for their use in other settings). Therefore, we constructed three IATs that aimed to compare the concept of 'violence' to the concepts of 'valence', 'hedonic value' and 'arousal' in turn. We tested these in a population of offenders where many had a strong history of violence (defined as a single act of extreme violence, or a long history of multiple but less serious violence, or both). First, we aimed to see if such tests produce reliable scores for an individual via internal reliability. Second, we aimed to see if scores on these IATs were related to measures of violence and criminality.

Typologies of violence

The motivations and reasons behind acts of violence are multifarious and have led to a range of taxonomies for offenders and violence. From a criminological point of view, there appears to be a small group of individuals (perhaps around 5%) with persistent antisocial behaviour which are responsible for over 50% of all known crimes (Farrington, Ohlin, & Wilson, 1986) and 50-70% of officially recorded violent crimes (Odgers et al., 2007; Odgers et al., 2008). These antisocial individuals tend to have elevated rates of negative personality traits such as hostility and negative affect, and this group are also associated with poor physical health and poor early social conditions (Odgers, 2009). Crucially, these individuals' psychological problems interact with their criminogenic environments to produce greater antisocial and pro-criminal attitudes in comparison to other offenders (Moffit, 1993; Paciello, Fida, Tramontano, Lupinetti, & Caprara, 2008).

From a psychodynamic point of view, Glasser (1996) draws a distinction between 'self-preservative' violence and 'malicious' or 'sadistic' violence. He describes self-preservative violence as reflecting an automatic response to danger and its aim is the elimination of a threat. The violent act is typically carried out in a state of high anxiety and is not particularly associated with other acts of violence or criminality. There is little regard for the victim and the perpetrator is consumed with the desire to be rid of a perceived threat to himself or his physical or emotional well-being. Some impulsive acts of homicide typify this type of violence. On the other hand, sadistic violence has as its aim the infliction of physical or emotional pain, which affords the perpetrator a degree of gratification. Fonagy and Target (1999) suggest that sadistic violence includes violent acts which are apparently carried out in a cold and callous state of mind and the violent act may even be enjoyed.

Thus, these two bodies of literature (the criminological and the psychodynamic) seem to converge on a general idea that there is a group of offenders that have persistent antisocial tendencies, have criminal mind sets and who may hold endemic pro-criminal attitudes, including positive views of violence and aggression. On the other hand, it is theorised that there are a group of violent offenders that do not hold these attitudes and who are violent despite of this. Violence from this group appears more likely to be reactive or defensive in nature. Hence, various researchers have suggested that murderers and non-murderers may have differences in their motivations (see Richetin & Richardson, 2008). For example, Clarke, McCarthy, Huband, Davies, Hollin, and Duggan (2016) compared mentally disordered offenders who had committed homicide to mentally disordered offenders who had committed other acts of violence. They found that homicide perpetrators were less likely to have suffered childhood adversity, had less generalised criminality, and were more likely to have a diagnosis of schizophrenia. This reinforces Glasser's (1996) typology of self-

perseverative violence and suggests that there was less personality pathology amongst the homicide perpetrators. Sherretts, Boduszek, Debowska, and Willmott (2017) compared homicide offenders to other groups of offenders on such characteristics as criminal social identity and psychopathic features. They found that the homicide offenders showed lower scores on measures of criminal cognitions and on ties to other criminal groups, supporting the notion that non-homicide offenders have more developed criminogenic features and criminal cognitions.

These different types of offender might have different violence-related cognitions that serve to fuel their violent actions (Richetin & Richardson, 2008). To approximate this distinction, we split our sample into those that have committed murder or attempted murder (homicide group) and those with no conviction for causing someone's death deliberately (non-homicide). Following previous research, outlined above, we predicted that the latter group (non-homicide) would have greater levels of convictions (for all offences), more convictions for violent offences, and that, crucially, their criminal/violent behaviours would be reflected (and most probably underpinned) in their implicit attitudes to violence. For the homicide-group (our proxy for the less criminogenic group of offenders in this high-risk offender sample) we hypothesised that their criminal/violent behaviour does not stem from pro-criminal attitudes and hence we did not predict any relationship between measures of past convictions and implicit attitudes to violence in this group. We acknowledge that the distinction between homicide vs. non-homicide offenders is only a crude proxy for the distinction proposed by Moffit (1993) for the more criminogenic and life-time anti-social offender and for the distinction outlined by Glasser (1996), but note that such a distinction that has been used in other empirical examinations of criminal motivations (Gray et al, 2003; Clarke et al, 2016; Sherretts et al., 2017).

Hypotheses and Research Purposes

We hypothesised that:

- 1) The implicit association test can be used to get a reliable measure of people's implicit cognitions relating to violence as "good", "enjoyable" and "exciting".
- 2) Increased levels of implicit attitudes that find violence either good, enjoyable, and/or exciting would be related to increased levels of self-reported violence and to criminal activity
- 3) This relationship would be stronger in those without a conviction for homicide (as it is this group of offenders who usually have the stronger criminogenic factors), while those with a conviction for homicide would not have any specific positive attitudes towards violence in any of the three domains (indeed, they may have negative attitudes to violence due to remorse for their actions).

Method

The research was conducted in HMP *****, a category B prison in the UK that is run as a therapeutic community for offenders diagnosed with a severe Personality Disorder. All experimental protocols and data collection methods were given ethical permission by both ***** Research and Advisory Committee and the NISCHR ***** Research Ethics Committee. All participants gave written informed consent to participate in the procedures.

Participants

All offenders were adult male offenders who had been admitted to the assessment unit at HMP ***** and, therefore, had a diagnosis of Personality Disorder. Consecutive admissions were approached between November 2012 and November 2014.

Our initial aim was recruit 100 participants, such that each group (homicide vs non-homicide) would consist of $N = 50$, which would give us a power of 0.80 to detect a moderate effect size ($r = 0.30$) at an alpha of 0.05 (one-tailed). However, our assignment to groups was based on criminal history so we could not control this assignment. We approached 111 possible participants and 110 (99%) agreed to participate. One participant then declined due to low levels of literacy and one withdrew on the basis that he was withdrawing from any non-compulsory research associated with the prison. This left a sample of 108 participants with usable data sets. As our IATs involve the speeded classification of words, we screened the sample for low verbal IQ using the Wechsler Test of Adult Reading (WTAR; Wechsler, 2001). All participants fell above the cut off score of 70. Average IQ was 98.0 (SD = 12.05, range 70 - 119). The average age of participants was 40 (SD=11.11) with a range from 23 to 64. Most participants described themselves as white (86; 79.63%), followed by black or mixed race (17; 15.74%), Asian (4; 3.7%) and finally, 'other' (1; 0.93%). The average number of convictions for participants was 16.46 (SD=15.79) and 46 of the participants had a conviction for homicide.

Stimuli and Materials

Implicit association tests. Three versions of the IAT were developed to test implicit associations and violence. The three IATs tested implicit associations between violence and valence, hedonic value, and arousal. All experiments were administered with a Microsoft Windows laptop using Direct RT software.

The concepts of violence and peaceful was represented via 12 pictures taken from the International Affective Picture System (IAPS – Lang et al, 1997). Six were chosen for their violent content (IAPS: 1525, 6244, 6313, 6315, 6360, 6350) and six for their peaceful content (IAPS: 1710, 1750, 2160, 2311, 2501, 5780) and were used in all three IATs. The picture set contained a range of elements such as animals, people, etc. to minimize any recoding effects (Rothermund, Teige-Mocigemba, Gast, & Wentura, 2009). In pilot testing there was unanimous agreement as to which category each picture represented. All pictures were 10-15 by 10-12 cm and were full colour.

The valence-IAT was used by representing the concepts of ‘good vs bad’ with 12 words, six of which were to be classified as ‘good’ (good, nice, healthy, happy, holiday, and flower) and six as ‘bad’ (vomit, cancer, poison, sick, sting, and cry) and had been so classified unanimously in pilot experiments. The hedonic value-IAT was used by representing the concepts of ‘enjoy vs dislike’ with 12 words, six of which were to be classified as ‘enjoy’ (enjoy, pleasure, delight, tasty, favourite, adore) and six as ‘dislike’ (hate, dislike, disgust, annoy, horrible, worst) and had been so classified unanimously in pilot experiments. The arousal-IAT was used by representing the concepts of ‘exciting vs boring’ with 12 words, six of which were to be classified as ‘exciting’ (exciting, fast, arouse, fire, buzz, electrify) and eight as ‘boring’ (calm, dull, slow, quiet, boring, yawn) and had been so classified unanimously in pilot experiments. All word stimuli were presented in Times New Roman text with approximate height of 0.6 cm.

Each IAT commenced with the presentation of an instruction screen. A reminder of the task requirements, i.e. how each stimulus was to be classified, was also present throughout the experiment on the screen (see Figure 1).

Each trial consisted of a fixation mark (cross) in the middle of the screen for 500 ms. This was then replaced by a stimulus (a picture or a word) in the middle of the screen and the

participant attempted to classify this stimulus as quickly as possible according to the instructions. The stimulus remained until the participant responded. The participant's response caused the screen to go blank and the next trial to commence. No feedback was given as to the correctness of the response.

All tasks consisted of four blocks. For the valence-IAT, in Block 1 the participant practised the task with the pairing of violence-good for 8 trials. Block 2 was identical to Block 1 but continued for 48 trials. In Block 3 the contingencies were reversed, and the participant practiced violence-bad for 8 trials. Block 4 consisted of the 48 trials with the same contingencies as Block 3. The hedonic value-IAT and arousal-IAT had the same structure, with violence-enjoy and violence-exciting contingencies for blocks 1 and 2, and violence-dislike and violence-boring in blocks 3 and 4. This streamlined IAT (with 4 blocks) differs from what might be regarded as a standard-IAT where each data collection block is preceded with blocks where each classification is presented in isolation. This was done to minimise fatigue and boredom in this offender population, especially given that several IATs were to be administered in one session. We have previously shown that such a streamlined IAT is as effective as the standard version (Brown, 2009) and have used this successfully in several settings (e.g., Brown, Gray, & Snowden, 2009; Snowden, Wichter, & Gray, 2008), as have others (Teachman, Gregg & Woody, 2001).

For the IAT tasks participants were administered the violence IAT trials in the following set order; valence, hedonic value, and arousal. This set order was used as we were interested in individual differences on the tasks rather than comparisons between tasks.

Proactive and reactive aggression questionnaire (RPQ). The RPQ (Raine et al., 2006) is a 23-item self-report questionnaire where the participant rates how often an aggressive behaviour has occurred in the past on a 3-point scale ('never', 'sometimes', 'often'). As well as providing a total aggression score, it provides two separate measures

relating to amounts of Proactive Aggression (when the person has been aggressive in a deliberate and planned manner; e.g., ‘had fights with others to show who was on top’), and Reactive Aggression (when the person has been aggressive in reaction to a particular circumstance or in an unplanned manner; e.g., ‘reacted angrily when provoked by others’). In our sample, we found strong correlations between the two scales ($r = 0.82$) and both scales behaved the same with respect to our independent variables. We, therefore, only present the total RPQ score in this paper. The measure has proven validity and reliability (Fossati et al., 2009). It showed excellent reliability in the present sample ($\alpha = .95$).

Previous convictions. Data about previous convictions were collected from the Offender Group Reconviction Scale (OGRS) and from the Offender Assessment System (OASYS). Only convictions were recorded. This had already been completed by the UK Probation Service as part of sentencing and management procedures. These convictions had been classified into 14 categories on the Offender Assessment System. We simplified this categorization system for our purposes. We classified the categories of murder, attempted murder, manslaughter, rape, actual bodily harm, grievous bodily harm, wounding, assault, and battery as violent crimes, and all others as non-violent crimes. The convictions were also classified as to whether they were convicted as a juvenile (under the age of 18). Finally, we also formed a “homicide group” which consisted of any offenders with a conviction for murder or attempted murder.

Data Analysis

Implicit Association Tests. The raw reaction times (RTs) and errors were transformed to produce a *D*-score using an approximation to the *D*₆ scoring technique (Greenwald, Nosek, & Banaji, 2003). RTs less than 300 ms or greater than 3000 ms were removed. Trials on which an error was made were punished by adding 600 ms to the RT for that trial. The IAT effect (the difference in performance between the incongruent and

congruent trials) was calculated as the difference in mean RTs for the blocks divided by the pooled standard deviation. Higher scores are interpreted as greater valence, hedonic value, and arousal for each of the three IATs.

For each IAT in turn, we removed the data for any participant who had excessive errors (> 30%), which resulted in the loss of two participants for the valence-IAT, three for the hedonic value-IAT and two for the arousal-IAT.

Results

Effects of age and IQ

To examine possible artefacts, we examined the effects of age and IQ on our independent and dependent variables. Age and IQ were not related to the criminological variables save that, as expected, increasing age was associated with a greater number of adult convictions with ($r = .29, p = .002$). Self-reported aggression was not associated with IQ but showed a negative relationship to age ($r = -.28, p = .008$). Age and IQ were not associated with any of the IAT scores ($ps > .14$).

For some statistical analyses the offenders were split into homicide vs non-homicide groups. The two groups did not differ in terms of age (36.5 vs 38.4 years, $p = .35$) or IQ (96.3 vs 100.2, $p = .09$).

Indices of Aggression.

Previous convictions – For all these variables the data showed high values of skew and kurtosis that were not corrected by data transformations. Therefore, we chose to use non-parametric statistics for these data.

Scores are illustrated in Table 1. As expected in this population of personality disordered offenders, there were high rates of all forms of conviction. We compared the number of convictions for the homicide and non-homicide offenders. As predicted, the non-

homicide group had a greater number of convictions than the homicide group ($Mdn = 16.5$ vs 5.5 ; Mann Whitney U, $p < .01$), but did not differ on number of violent convictions ($Mdn = 2.0$ vs 2.0).

RPQ - Nineteen participants refused to complete the Reactive and Proactive Aggression scales, leaving 89 suitable cases.

Scores are illustrated in Table 1. A comparison of the non-homicide and homicide groups did not show any significant differences in self-reported aggression ($M = 20.7$ vs 18.6).

The various measures of antisocial behaviour and violence were correlated with each other and the results are shown in Table 1. The measures of offending behaviour (convictions, violent convictions, juvenile convictions and adult convictions) showed strong correlations (as expected as there is actual overlap between most of these categories) of $.34$ - $.90$ but were only moderately related to self-reported violence ($r_s = .17$ to $.28$).

Implicit Association Tests

Valence-IAT. To examine the reliability of this task (see hypothesis 1), we split the trials into odd and even trials and calculated two D-scores based on the odd and the even trials for each participant. We then correlated these two scores and applied the Spearman-Brown formulation to correct for the loss of data due to splitting. The resulting reliability correlation was very good ($r = .81$, $p < 001$). The mean for the valence-IAT was negative ($M = -0.77$, $SD = 0.63$) and was significantly different from zero, $t(101) = -12.35$, $p < .001$, $d = 2.46$, which indicates that the sample associated violence with the concept of 'bad' (or peaceful with 'good'). The scores for the homicide and non-homicide groups did not differ (-0.88 vs -0.68 , $p = .18$).

Hedonic value-IAT. Reliability was calculated as before. The resulting reliability correlation was good ($r = .77$, $p < 001$). The mean for the hedonic value-IAT was negative

($M = -0.63$, $SD = 0.59$) and was significantly different from zero, $t(101) = -10.67$, $p < .001$, $d = 2.21$, which indicates that the sample associated violence with the concept of ‘dislike’ (or peaceful with ‘enjoy’). The scores for the homicide and non-homicide groups did not differ (-0.64 vs -0.63 , $p = .98$).

Arousal-IAT. Reliability was calculated as before. The resulting reliability correlation was only moderate ($r = .56$, $p < .001$). The mean for the arousal-IAT was positive ($M = 0.15$, $SD = 0.39$) and was significantly different from zero, $t(102) = 3.81$, $p < .001$, $d = 0.75$, which indicates that the sample associated violence with the concept of ‘excitement’ (or peaceful with ‘boredom’). The scores for the homicide and non-homicide groups did not differ (0.15 vs 0.16 , $p = .93$).

Attitudes to Violence.

Hypotheses two and three were that higher scores on the IATs would be associated with greater levels of self-reported aggression and official records of criminal activity, and this would be greater for those in the non-homicide group. Therefore, data from the IATs was correlated against the measures of violence and the results are shown in Table 2.

Valence-IAT. The conventional valence-IAT was not significantly correlated with any of the measures of violence or antisocial behavior in the total sample, or when the samples were split into homicide and non-homicide groups.

Hedonic value-IAT. The hedonic value-IAT score showed significant positive correlations with number of previous convictions, convictions for juvenile offences, and a trend for significance with number of adult convictions. However, it was not significantly related to self-report measures of violence. When the participants were split into homicide and non-homicide groups, significant correlations emerged in the non-homicide group, as predicted, but not in the homicide group. Statistical analysis (Fisher r -to- z transformation: p

< .05) showed that the correlations for number of juvenile convictions were significantly greater in the non-homicide group than the homicide group.

Arousal-IAT. The arousal-IAT score showed little relationship to the conviction measures for the sample, although there was a trend towards a negative relationship with the measure of juvenile convictions. However, when the participants were split into homicide and non-homicide groups, the non-homicide group showed a positive relationship between the arousal-IAT and self-reported aggression on the RPQ.

Discussion

We developed three versions of the implicit association test (IAT) to measure people's automatic associations to stimuli depicting violent actions. These three IATs aimed to separate different components of violence-related cognitions, namely those related to valence, hedonic-value and arousal. We reasoned that the dimensions of 'enjoyment' and of 'excitement' might be good predictors of violent actions, possibly better than a judgement of the valence of violence (good versus bad) and that the development of such implicit measures might allow practitioners a valuable dynamic instrument for assessing violence-related cognitions.

Our first hypothesis was that the IAT would be a reliable measure of violence-related cognitions. This was achieved for the valence-IAT and the hedonic value-IAT, but the arousal-IAT did not reach acceptable levels of reliability. We discuss reasons for this below.

Our second hypothesis was that these IATs would be associated with increased levels of violence and criminal activity. For the total sample, we found only little evidence for the validity of any of the violence-IATs. However, our third hypothesis predicted that these relationships would be greater in the non-homicide offenders (who were hypothesised to have more pro-criminal and antisocial attitudes) and, as expected, had greater number of convictions in the present sample than in the homicide offenders. The results showed

tentative support for this hypothesis and show that a consideration of the heterogeneity of offenders is important.

Is the IAT a Valid Measure?

Are the low correlations between the IAT scores and measures of violence real or an artefact of our experimental methods? It might be argued that indirect techniques, including the IAT, are not able to measure people's attitudes to violence. Many indirect measures of attitudes have poor psychometric properties, including reliability (Cunningham, Preacher, & Banaji, 2001). However, the IAT was chosen as our preferred technique as it has been shown to have quite good psychometric properties (Bar-Anan & Nosek, 2014). This was confirmed in the present studies, where the reliability coefficients for the valence-IAT and the hedonic value-IAT were very good ($> .75$). However, the reliability for the arousal-IAT was low and may account for the deficient performance of this IAT with respect to predicting violent behaviours.

The weaker reliability of the arousal-IAT might arise from one of several sources. First, when constructing the arousal-IAT we had to consider what the opposite of 'excitement' was (which we decided was the concept of 'boring'). However, it seems unlikely to us that many people would ever associate violence with being boring. Most people are likely to become highly aroused when seeing or committing violence, but for some individuals this may be due to fear, whilst for others it might be due to excitement. Hence, we might expect less inter-individual variation on the arousal-IAT (it is notable that the SD for this IAT was considerably less than for the other IATs) and hence the weaker estimates of reliability. Second, it may be difficult to capture the concept of 'arousal' or 'excitement' using only semantic associations. It seems likely that this association to violence may be better measured via physiological arousal such as skin conductance or heart rate responses rather than relying on cognitive responses.

If the relative lack of correlation between our IATs and the measures of violence is not due to the poor psychometric properties of the IATs, could it be due to the poor measurement of violence? Measuring violence is difficult and so we relied on two very different sources of information. First, we took a self-report measure of past violence (RPQ). The problems of self-report have been rehearsed elsewhere (e. g., Paulhus, Bruce, & Trapnell, 1995). It seems likely that some participants may have deliberately misrepresented themselves to either appear 'good' (i.e., less violent) or to appear 'bad' (i.e., more violent) in such a prison setting, despite the assured anonymity of their responses. It is also probable that such offenders 'normalise' themselves with regard to their offender peers rather than to the rest of society via self-referencing effects. These participants were part of a therapeutic community and were therefore exposed to discussions of each other's offending behavior.

Second, we took measures of criminal activity via official records. Clearly, such measures are limited by the underlying offending behavior being identified by authorities and a conviction being secured. It is known that most crimes, including violent crimes, do not result in a conviction. Further, time spent incarcerated can lead to offenders having less time available to commit other crimes, thus skewing behavioral measures of violence. Hence, these variables (that include antisocial acts as well as violent acts) are at best a proxy for actual violence or violence propensity.

Both methods (self-report and official records), therefore have limitations and it is perhaps not surprising that our data show that there are only modest correlations between the two domains (see Table 1). Future studies are needed to consider if more positive results occur in populations with less violent histories who may more honestly self-report their past history of proactive and reactive aggression, or in experiments where violent behaviour could actually be observed, such as in laboratory settings (e.g., Anderson, Lindsay, & Bushman, 1999).

Implicit Attitudes to Violence.

We found that the hedonic-IAT was correlated with measures of general criminality (as measured by overall number of convictions and juvenile convictions), and that this was most apparent in the non-homicide group (with the correlation with juvenile convictions being statistically different between the two groups). We had predicted that the hedonic value-IAT would be predictive of such antisocial acts in this group and so this finding supports this hypothesis (although the lack of a correlation to the number of violent convictions does not). The finding that the hedonic value-IAT was significantly related to behavioral measures of criminal behaviour, whereas the valence-IAT was not, may suggest that this change in emphasis from whether someone regards violence as good vs bad, to whether they enjoy or dislike it may be a fruitful direction for further development of indirect measures of violence-related cognitions.

We also found that the arousal-IAT was associated with self-reported aggression in the non-homicide group. We interpret this result to imply that some individuals are aroused or excited by aggression or violence and this may underpin their violent behaviour. However, this interpretation is not supported by the data from the conviction data (overall or violent). Increased levels of arousal associated with self-reported violence on the arousal-IAT were not associated within increasing levels of convictions. The difference in the pattern of results from the self-report measure and the official records may reflect differences in willingness to report aggression. It maybe that people who are automatically excited by violence are more willing to report aggressive acts on self-report measures.

Strengths and Limitations

The major strengths of the study are that we have examined three novel methods of examining implicit cognitions about violence, and we have done so in a population in which there is a strong history of serious violent behaviour.

We have already outlined some of the major limitations concerning both the IAT measures and the measures of criminality, and self-reported violence available to us in this sample. Here, we consider other possible limitations.

First, our data analysis strategy involved many comparisons and we did not attempt to adjust alpha levels to accommodate this. We justify this in terms of this being an exploratory analysis and the first attempt to use these novel measures to explore attitudes to violence in a high-risk offender sample but recognise that in doing so we increase the chance that some (or all) of the significant results we report may be Type I errors. If this body of research is developed further, then more stringent tests will need to be applied. We should also acknowledge that our sample size was designed to detect moderate effect sizes and so we may have committed some Type II errors in the failure to detect smaller effects.

Second, the sample we used included participants with severe antisocial histories, nearly all of which contain serious interpersonal violence. We might have, therefore, set our IATs a very stringent test of attempting to detect individual differences within a rather homogenous group. Studies that compare violent offenders to other non-violent groups may meet with greater success.

Finally, our grouping of the offenders based solely on a conviction for homicide (or not) is crude. As before, we did this as a first exploratory analysis into whether such a division was worthwhile empirically. More detailed analyses of offenders' criminal and violent behaviours might produce a more refined and accurate classification of offender typology.

Research Implications

The present research leaves many unanswered questions. For example, the IATs we developed are only three among the many that might be thought to relate to violence-related cognitions. For example, IATs could be constructed that examine associations between

violence and other emotions, such as fear, anxiousness, and anger. Further, other indirect measures, such as a priming task (see Snowden et al., 2008) may prove useful in examining reactions to specific exemplars of violence rather than the abstract concept of violence (e.g., specific cognitions about gang-related violence, or to specific victim groups (hate-crimes, terrorist offences)).

Clinical and Policy Implications

The over-arching aim of our research programme was to find techniques for measuring violence-related cognitions in offenders who might lack insight into these underlying cognitions and/or may be unwilling to express them honestly. We hope to be able to use these implicit techniques to identify treatment needs and to track any therapeutic change. This study was the first to use indirect methods to examine violence-related cognitions as enjoyable or exciting, but the results were not strong. We are certainly not yet in a position where these IATs could be used for clinical or forensic assessments of pro-criminal attitudes. However, we hope that the few positive results, and our critique of our methods, might inspire further work to refine and improve on the potential of such indirect techniques to measure violence-related cognitions.

References

- Anderson, C. A., Lindsay, A. J., & Bushman, B. J. (1999). Research in the psychological laboratory: Truth or triviality? *Current Directions in Psychological Science*, 8, 3-9.
- Bar-Anan, Y., & Nosek, B. A. (2014). A comparative investigation of seven indirect attitude measures. *Behavior Research Methods*, 46, 668-688.
- Bluemke, M., Crombach, A., Hecker, T., Schalinski, I., Elbert, T., & Weierstall, R. (2017). Is the implicit association test for aggressive attitudes a measure for attraction to violence or traumatization? *Zeitschrift Fur Psychologie-Journal of Psychology*, 225(1), 54-63. doi: 10.1027/2151-2604/a000281
- Bowes, N. & McMurrin, M. (2013) Cognitions supportive of violence and violent behavior, *Aggression and Violent Behavior*, 18, 660-665.
- Brown, A. S. (2009). *Developing an Implicit Association Test for forensic use: Discriminating paedophiles from other offenders*. (PhD), Cardiff University, UK.
- Brown, A. S., Gray, N. S., & Snowden, R. J. (2009). Implicit measurement of sexual associations in child sex abusers: Role of victim type and denial. *Sexual Abuse-A Journal of Research and Treatment*, 21, 166-180.
- Buss, A. H., & Perry, M. (1992). The aggression questionnaire. *Journal of Personality and Social Psychology*, 63, 452-459.
- Chassin, L., Presson, C. C., Sherman, S. J., & Edwards, D. A. (1991). Four pathways to young-adult smoking status - adolescent social-psychological antecedents in a midwestern community sample. *Health Psychology*, 10, 409-418.
- Clarke, M., McCarthy, L., Huband, N., Davies, S., Hollin, C., & Duggan, C. (2016). The Characteristics and course after discharge of mentally disordered homicide and non-homicide offenders. *Homicide Studies*, 20, 80-97.

- Cunningham, W. A., Preacher, K. J., & Banaji, M. R. (2001). Implicit attitude measures: consistency, stability, and convergent validity. *Psychological Science, 12*, 163-170.
- Eckhardt, C. I., Samper, R., Suhr, L., & Holtzworth-Munroe, A. (2012). Implicit attitudes toward violence among male perpetrators of intimate partner violence: A preliminary investigation. *Journal of Interpersonal Violence, 27*, 471-491.
- Farrington, D., Ohlin, L., & Wilson, J. Q. (1986). *Understanding and Controlling Crime*. New York: Springer-Verlag.
- Fonagy, P., & Target, M. (1999). Towards understanding violence: The use of the body and the role of the father. . In R. J. Perelberg (Ed.), *Psychoanalytic Understanding of Violence and Suicide* (pp. 51 - 72). London: Routledge.
- Fossati, A., Raine, A., Borroni, S., Bizzozero, A., Volpi, E., Santalucia, I., & Maffei, C. (2009). A Cross-cultural study of the psychometric properties of the Reactive-Proactive Aggression Questionnaire among Italian nonclinical adolescents. *Psychological Assessment, 21*, 131-135.
- Friese, M., Wanke, M., & Plessner, H. (2006). Implicit consumer preferences and their influence on product choice. *Psychology & Marketing, 23*, 727-740.
- Glasser, M. (1996). The assessment and management of dangerousness: the psychoanalytical contribution. *Journal of Forensic Psychiatry, 7*, 271-283.
- Gray, N. S., MacCulloch, M. J., Smith, J., Morris, M., & Snowden, R. J. (2003). Violence viewed by psychopathic murderers. *Nature, 423*, 497-498.
- Greenwald, A. G., McGhee, J. L., & Schwartz, J. L. (1998). Measuring individual differences in implicit cognition: The implicit association test. *Journal of Personality and Social Psychology, 74*, 1464-1480.

- Greenwald, A. G., Nosek, B. A., & Banaji, M. R. (2003). Understanding and using the implicit association test: I. An improved scoring algorithm. *Journal of Personality and Social Psychology*, *85*, 197-216.
- Huijding, J., de Jong, P. J., Wiers, R. W., & Verkppeijen, K. (2005). Implicit and explicit attitudes toward smoking in a smoking and non-smoking setting. *Addictive Behaviors*, *30*, 949-961.
- Moffitt, T. E. (1993) Adolescence-limited and life-course persistent antisocial behavior: A developmental taxonomy. *Psychological Review*, *100*, 674-701.
- Odgers, C. L. (2009). The life-course persistent pathway of antisocial behaviour: risks for violence and poor physical health. In S. Hodgins, E. Viding, & A. Plodowski (Eds.), *The Neurobiological Basis of Violence: Science and Rehabilitation* (pp. 23 - 41). Oxford, UK: Oxford University Press.
- Odgers, C. L., Caspi, A., Broadbent, J. M., Dickson, N., Hancox, R. J., Harrington, H., . . . Moffitt, T. E. (2007). Prediction of differential adult health burden by conduct problem subtypes in males. *Archives of General Psychiatry*, *64*(4), 476-484.
- Odgers, C. L., Moffitt, T. E., Broadbent, J. M., Dickson, N., Hancox, R. J., Harrington, H., . . . Caspi, A. (2008). Female and male antisocial trajectories: From childhood origins to adult outcomes. *Development and Psychopathology*, *20*(2), 673-716.
- Office for National Statistics (2016). *Crime in England and Wales: Year ending September 2015*. crimestatistics@ons.gsi.gov.uk: Home Office UK.
- Olivera-La Rosa, A., Villacampa, J., Amador, O., Corridi, G., Munar, E., Acosta, S., & Rossello, J. (in press). Implicit attitudes toward violence in a sample of adolescent offenders with conduct disorder. *Journal of Interpersonal Violence*.

- Ostrosky-Solis, F., Rebollar, C. R., Garcia, A. V., & Villalpando, R. O. (2009). Implicit associations as a method for detection of violence and psychopathy. *Revista Chilena de Neuropsicología, 4*, 170-178.
- Paciello, M., Fida, R., Tramontano, C., Lupinetti, C., & Caprara, G. V. (2008). Stability and change of moral disengagement and its impact on aggression and violence in late adolescence. *Child Development, 79*, 1288–1309.
- Paulhus, D. L., Bruce, M. N., & Trapnell, P. D. (1995). Effects of self-presentation strategies on personality profiles and their structure. *Personality and Social Psychology Bulletin, 21*, 100-108.
- Perugini, M., Richetin, J., & Zogmaister, C. (2010). Prediction of behavior. In B. Gawronski & B. K. Payne (Eds.), *Handbook of implicit social cognitions* (pp. 255 - 277). London: The Guildford Press.
- d.
- Polaschek, D. L. L., Calvert, S. W., & Gannon, T. A. (2009). Linking violent thinking implicit theory-based research with violent offenders. *Journal of Interpersonal Violence, 24*, 75-96.
- Raine, A., Dodge, K., Loeber, R., Gatzke-Kopp, L., Lynam, D., Reynolds, C., . . . Liu, J. H. (2006). The reactive-proactive aggression questionnaire: Differential correlates of reactive and proactive aggression in adolescent boys. *Aggressive Behavior, 32*, 159-171. doi: 10.1002/ab.20115
- Richetin, J., & Richardson, D. S. (2008). Automatic processes and individual differences in aggressive behavior. *Aggression and Violent Behavior, 13*, 423-430.
- Robinson, T. E., & Berridge, K. C. (2003). Addiction. *Annual Review of Psychology, 54*, 25-53.

- Rothermund, K., Teige-Mocigemba, S., Gast, A., & Wentura, D. (2009). Minimizing the influence of recoding in the Implicit Association Test: The Recoding-Free Implicit Association Test (IAT-RF). *Quarterly Journal of Experimental Psychology*, *62*, 84-98.
- Snowden, R. J., & Gray, N. S. (2010). Implicit social cognition in forensic settings. In B. Gawronski & B. K. Payne (Eds.), *Handbook of implicit social cognitions* (pp. 522-534). London: The Guildford Press.
- Snowden, R. J., MacCulloch, M. J., Smith, J., Morris, M., & Gray, N. S. (2004). Implicit affective associations to violence in psychopathic murderers. *Journal of Forensic Psychiatry and Psychology*, *15*, 620-641.
- Snowden, R. J., Wichter, J., & Gray, N. S. (2008). Implicit and explicit measurements of sexual preference in gay and heterosexual men: A comparison of priming techniques and the implicit association task. *Archives of Sexual Behavior*, *37*, 558-565.
- Teachman, B. A., Gregg, A. P., & Woody, S. R. (2001). Implicit associations for fear-relevant stimuli among individuals with snake and spider fears. *Journal of Abnormal Psychology*, *110*, 226-235.
- Tibboel, H., De Houwer, J., & Van Bockstaele, B. (2015). Implicit measures of 'wanting' and 'liking' in humans. *Neuroscience and Biobehavioral Reviews*, *57*, 350-364. 15
- Wechsler, D. (2001). *Wechsler test of adult reading (WTAR-UK)*: Harcourt Assessment.
- Weldon, S., & Gilchrist, E. (2012). Implicit theories in intimate partner violence offenders. *Journal of Family Violence*, *27*, 761-772.
- Wilson, T. D., & Dunn, E. W. (2004). Self-knowledge: Its limits, value, and potential for improvement. *Annual Review of Psychology*, *55*, 493-518.
- Zwets, A. J., Hornsveld, R. H. J., Muris, P., Huijding, J., Kanters, T., Snowden, R. J., & van Marle, H. (2015). Implicit attitudes toward violence and their relation to psychopathy,

aggression, and socially adaptive behaviors in forensic psychiatric inpatients. *Journal of Forensic Psychiatry & Psychology*, 26, 632-651.

Table 1. *Descriptive statistics and correlations (Spearman's rho) between the various measures of offending and violence.*

	n	Mean or median	SD or Quartiles	1	2	3	4
1. RPQ	89	19.9	5.65	-	.		
2. Convictions	108	12.0	4 - 24	.28**	-		
3. Violent convictions	108	2.0	1 - 4	.26**	.46**	-	
4. Juvenile convictions	108	3.0	0.25 - 8	.28**	.67**	.38**	-
5. Adult convictions	108	6.5	2 - 16	.17+	.90**	.42**	.34**

+ $p < .10$, * $p < .05$, ** $p < .01$

Table 2. *Correlations (Spearman’s rho) between the IATs and measures of offending and violence.*

	RPQ	Convictions	Violent convictions	Juvenile convictions	Adult convictions
<i>Valence-IAT</i>					
All	-.00 ^a	-.14 ^b	-.14 ^b	-.04 ^c	-.16 ^c
Homicide	-.06 ^f	-.13 ^g	-.15 ^g	-.08 ^g	-.15 ^g
Non-homicide	.10 ^j	-.10 ^k	-.15 ^k	.01 ^k	-.07 ^k
<i>Hedonism-IAT</i>					
All	.07 ^a	.18* ^b	.02 ^b	.18* ^c	.16 ⁺ ^c
Homicide	.01 ^f	.07 ^g	.09 ^g	-.03 ^g	.14 ^g
Non-homicide	.17 ^j	.29* ^k	-.05 ^k	.40** ^k	.23 ⁺ ^k
<i>Arousal-IAT</i>					
All	.08 ^d	-.08 ^e	-.14 ^e	-.18* ^b	-.01 ^b
Homicide	.03 ^h	-.19 ⁱ	-.11 ⁱ	-.29* ⁱ	-.11 ⁱ
Non-homicide	.29* ^l	.11 ^m	-.15 ^m	-.05 ^m	.15 ^m

+ $p < .10$, * $p < .05$, ** $p < .01$

Number of complete datasets: a = 81, b = 101, c = 100, d = 82, e = 102, f = 40, g = 53, h = 38, i = 51, j = 39, k = 47, l = 42, m = 50

Figure 1. Illustration of the appearance of a typical trial where a picture stimulus was presented (due to copyright issues related to IAPs, the picture used here is illustrative and was not used in actual experiment).

PEACE
or
EXCITING



VIOLENT
or
BORING