AN EXPLORATION OF INTIMATE PARTNER VIOLENCE ACROSS THE LIFESPAN AND THE EFFECTIVENESS OF TREATMENT

By

SAMANTHA GOSWELL

A thesis submitted to the Centre for Forensic and Criminological Psychology of
The University of Birmingham for the degree of ForenPsyD

Centre for Forensic and Criminological Psychology
School of Psychology
College of Life and Environmental Sciences
The University of Birmingham
October 2014

Abstract

This thesis aims to further academic understanding of intimate partner violence (IPV) treatment and wider practice issues across the lifespan. This is achieved by three pieces of work; a systematic literature review, an empirical study and a psychometric critique. The opening chapter in this thesis provides a background to the issues pertinent to IPV interventions among adolescents and adults. A systematic literature review is presented in Chapter two examining the association between attitudes condoning violence and adolescent IPV (AIPV) perpetration and/or victimisation. Significant yet modest associations were found although methodological differences impeded firm conclusions. Implications of the findings pertaining to gender differences, acceptability of AIPV-supportive attitudes, prevention and intervention are discussed. The empirical study in Chapter three investigates the effectiveness of an IPV treatment programme among a community sample of male perpetrators (N=259). Results show positive effect in relation to self-reported psychological change. The behavioural outcome measure of recidivism yielded more complex results. Non-completers were found to be higher risk and reported higher levels of avoidant attachment and jealousy. Recidivists were slightly younger in age. Chapter four provides an in-depth critique of Spielberger's (1999) STAXI-2; a standardised and widely utilised tool measuring an individual's assessed experience and expression of anger and one that was used to evaluate the empirical study. A discussion is presented regarding its' general psychometric properties. The utility of the STAXI-2 among IPV adult and AIPV adolescent samples is considered in context of the complex interplay between anger experience and expression in this paradigm. Chapter five draws the thesis together through summarising the findings and considering the implications in light of identified limitations. A discussion relating to future directions in research and practice is presented.

Acknowledgements

Many people have contributed to making this thesis possible. Firstly, I would like to thank my academic supervisor Professor Leam A. Craig for his advice, support, contributions and guidance. A big thank you is extended to Sue Hanson who has been so kind, supportive and encouraging. I would also like to express my gratitude to Professor Anthony. R. Beech for his encouraging and motivational words.

Thank you to my family, friends and colleagues for their support and understanding. I would like to thank; my mum and sister for their constant support, sound advice, encouragement and patience during times of difficulty; my sister's partner for helping me put things into perspective and my beautiful nephew, whose face can brighten up any day.

My biggest thanks have to be given to my Dad, Dave. Without his belief, support, wisdom and love, I would not have been able to start, or finish, this journey. Thank you for all that you do Dad, I love you.

Finally, thank you to those at Thames Valley Probation Trust for enabling the research to be conducted and also to the staff within the Thames Valley Programmes Team for their vital role in enabling the data collection.

Table of Contents

List of Tables

List of Figures

List of Appendices

Chapter One

Introduction to the Thesis

The phenomenon of Intimate Partner Violence (IPV) has been increasingly identified as a significant social problem since the 1970's. Within the UK, there is no statutory definition of IPV making it difficult for professionals to adopt consistent terminology that describes the complexity of violence within intimate relationships (Dixon & Graham-Kevan, 2010).

More recently, there have been marked developments in defining this phenomenon including the revision and extension of the UK cross-government definition of domestic violence and abuse. The definition now includes 16 and 17 year olds, so-called 'honour' based violence, female genital mutilation, forced marriage and incorporates controlling and coercive behaviours (Home Office, 2013). The latter is now classified in the Serious Crime Act 2015 as a criminal offence. The cross-government definition of domestic violence and abuse is:

'any incident or pattern of incidents of controlling, coercive, threatening behaviour, violence or abuse between those aged 16 or over who are or have been intimate partners or family members regardless of gender or sexuality. The abuse can encompass but is not limited to: psychological, physical, sexual, financial or emotions' (Home Office, 2013).

The definition also describes controlling behaviour as:

'a range of acts designed to make a person subordinate and/or dependent by isolating them from a course of support, exploiting their resources and capacities for personal gain, depriving them of the means needed for independence, resistance or escape and regulating their everyday behaviour' (Home Office, 2013). And coercive behaviour as: *'an act or a pattern of acts of assault, threats, humiliation and intimidation or other abuse that is used to harm, punish, or frighten their victim'* (Home Office, 2013).

The construct of IPV utilised by academics and practitioners has previously been criticised as generally poorly defined (Bowen, 2011). Within the academic literature, the term Intimate Partner Violence (IPV) has been largely accepted to encompass non-violent abusive behaviours and aggressive and violent behaviours though technical differences between these terms are recognised (Archer, 2000). For the purposes of this thesis, IPV is defined as "any form of aggression and/or controlling behaviours used against a current or past intimate partner of any gender or relationship status" (Dixon & Graham-Kevan, 2011, p. 1145). This can include physical, sexual or psychological aggression.

Despite the definitions recognising distinct forms of IPV, research and prevalence surveys continue to be focussed upon physical IPV (Dobash & Dobash, 2004). There remains a lack of consistent evidence on the prevalence and impact of psychological IPV and controlling behaviours across the lifespan (Leen et al., 2013; see Chapter two).

Violent and abusive behaviour within intimate partner relationships reportedly accounted for 15% of all violent incidents in England and Wales in 2011/2012 (Strickland, 2013). Furthermore, statistics indicate that 1.2 million women (7%) and 800,000 (5%) men were victims of IPV in the UK in 2011/12 (Office for National Statistics; ONS, 2013). The Crime Survey for England and Wales (CSEW) reported that approximately five million women (31%) and 2.9 million men (18%) had experienced IPV since the age of 16 amongst those aged 16-59 years old (ONS, 2013). These figures included the prevalence of sexual and non-sexual IPV. This highlights the disparity between reporting and conviction since most reported IPV does not result in a conviction (Hester, 2006).

The true figure of IPV is difficult to ascertain as IPV is considered a 'hidden crime'. This is associated with the reluctance of victim reporting incidents (Strickland, 2013; Tjarden & Thoennes, 2000) and continued lack of public awareness as to what constitutes IPV (Home Office, 2013). These figures highlight that IPV remains a social issue to be addressed

particularly since there has been no statistically significant decrease in the level of reported since the 2008/9 CSEW. That said; the current figures are lower than the 2004/5 CSEW.

IPV across the lifespan

IPV is not limited to adult-age populations. A survey was conducted by the NSPCC to examine the prevalence of IPV amongst adolescents aged between 13 and 16+ years within the UK amid evidence of an increase amongst younger age groups (Barter, McCarry, Berridge & Evans, 2009).

The prevalence rate of adolescent intimate partner violence (AIPV) perpetration is identified as being between 10% and 50% (Jackson, 1999; see Leen et al., 2013 for a review). Survey data obtained from 183 victims of AIPV (aged under 18 years; gender not specified) found that 78% of the victims were experiencing controlling behaviours such as threats to kill and/or exposed to sexual activity, isolation from others and being put down in public (CAADA, 2012). Additionally, 76% of the sample was identified as experiencing physical abuse and 53% were experiencing harassment and stalking behaviours (CAADA, 2012). Gender differences within the paradigm of AIPV indicate that girls may be perpetrating more than boys although girls are more likely to be injured (Leen et al., 2013; Swahn, Simon, Arias & Bossarte, 2008; Temple, Shorey, Tortolero, Wolfe & Stuart, 2013). Cultural attitudes and understanding of AIPV contributes to the difficulty in obtaining accurate prevalence data since adolescents do not necessarily view abusive behaviours within relationships as destructive (Schumacher & Smith Slep, 2004).

From a developmental theory perspective, partner violence is considered as an expression of continued anti-social behaviour across the lifespan (Ehrensaft, Moffitt & Caspi, 2004). Early-age violent delinquency was found to predict IPV at age 21 equally for male and female participants in the robust, and on-going, Dunedin Study (Moffitt, Caspi, Rutter &

Silva., 2001). Further, this study demonstrated the stability of anti-social behaviour from adolescence to adulthood with some, albeit few, exhibiting life-course persistent antisocial behaviour. O'Leary et al. (1989) found that those who engaged in IPV during the dating phase of their relationship were three times more likely to use IPV within their marriage however this referred to relationship status as opposed to the age of IPV onset. This highlights that IPV may be a life-long issue for some (Moffitt et al., 2001).

There has been much research examining the association between environmental factors in childhood and subsequent IPV. Lussier et al. (2009) conducted a robust longitudinal study and provided evidence that an anti-social familial environment in childhood increases the risk of IPV perpetration in adulthood. Studies have identified a link between witnessing IPV within the family home and its predictive and meditative role in adolescents' use of IPV (Narayan, Englund, Carlson & Egeland, 2014) and adults' use of IPV (Roberts, Gilman, Fitzmaurice, Decker & Koenen, 2010; Theobald & Farrington, 2012).

A meta-analysis by Stith et al. (2000) found a weak-moderate association for the intergenerational cycle of abuse, particularly amongst community samples. Stith et al. (2000) found stronger associations among clinical samples where such samples were considered to observe more severe forms of abuse. Many of those who witness IPV do not necessarily go on to perpetrate IPV however the relationship is complex and differential effects upon individuals must be considered. Indeed, clinical consequences can occur from non-clinically abusive behaviour, for example, hospital visits and psychological ill-health for victims (Ehrensaft et al., 2004).This further highlights that childhood, adolescent and adult ages within the paradigm of IPV should be considered from a life-course perspective.

With the above considered, the question of whether AIPV perpetration, particularly in early adolescence, acts as a precursor to adult IPV perpetration remains under-researched and inconclusive (Shorey, Cornelius & Bell, 2008). Some researchers have hypothesised this link,

commenting that behaviours in adolescent relationships may provide the 'script' to inform later behaviours within relationships (Wekerle & Wolfe, 1999; see Chapter two).

Some longitudinal research has found that IPV perpetration may decrease across the life-span in general (Shortt, Capaldi, Kim, Owen & Feingold, 2012) and indeed, a significant proportion of men who engage in less severe forms of IPV do desist (Walker, Bowen & Brown, 2013). However, examining each of the specific IPV categories yields more complex results. The use of psychological aggression among men in their early 20's predicted further psychological aggression ten to twelve years later and physical aggression predicted the same behaviour approximately seven years later (Shortt et al., 2012). This study also highlighted dyadic factors within a developmental systems approach when examining the stability of IPV perpetration and victimisation (Stith, McCollum, Amanor-Boadu & Smith, 2012). The stability of male IPV perpetration was associated with relationship transitions and the new partner's IPV perpetration (or lack of). Similar results have been found in examining adolescent dating couples, whereby IPV was more stable within continuous dating relationships over a one year period (Fritz & Slep, 2000). This has implications for the application of intervention programmes aimed at addressing IPV perpetration, particularly that of couples-based treatment (Stith et al., 2012) and mate-selection research (Moffitt et al., 2001).

Theories of IPV

Theoretical debates between the feminist perspective, which places the role of patriarchal attitudes as a central risk factor for IPV (Respect, 2008) and the gender inclusive perspective of IPV, which considers a broader interaction of factors, has been well documented (Bell & Naugle, 2008; Dixon & Graham-Kevan, 2011; Dutton & Corvo, 2006; Hamel, 2007, 2009; Straus, 2008).

The feminist perspective places IPV within a socio-cultural context based on the notion that IPV is "inextricably linked to attempts to dominate and control women" (Dobash & Dobash, 1979, as cited in Dobash, Dobash, Wilson & Daly, 1992, p.71) and is overwhelmingly perpetrated by males as a result of this (Dobash & Dobash, 2004). Feminist-based programmes are therefore centred on psycho-educational techniques aimed at re-educating men in relation to their beliefs towards women and relationships. The most utilised IPV treatment model currently in North America (Gondolf, 2007) and the UK (Hamilton, Koehler & Lösel, 2013) is the Duluth model (Pence & Paymar, 1993.) which came into existence as a means of providing a more structured and collaborative approach to rehabilitating IPV offenders as opposed to the more ad hoc counselling approaches of the 1970's and 80's (See Bowen, 2011, for a comprehensive account of IPV intervention development).

Offending behaviour intervention programmes are required to be guided by empirical findings and based on the accurate assessment of Risk, Need and Responsivity principles (RNR; Andrews & Bonta, 2010a; see Chapter four). However the development and evaluation of IPV perpetrator interventions has been negatively impacted by the gendered focus within the paradigm as discussed below (Bowen, 2011; Dixon et al., 2012).

Empirical evidence has increasingly questioned the efficacy of the socio-cultural associations to IPV (Archer 2000, 2006, Dixon & Graham-Kevan, 2011; Dutton & Corvo, 2006; Straus, 2011). Moreover, large scale survey research with national samples and empirical meta-analyses have consistently found similar prevalence rates for male and female perpetration (Archer, 2000, 2006; Dutton, 2007; Lussier, Farrington & Moffitt, 2009; Moffitt et al., 2001; Straus, 2008). It is argued that the debate largely stems from methodological sampling differences; non-clinical community cases of IPV versus clinical women's shelter samples (Ehrensaft, Moffitt & Caspi, 2004).

The feminist perspective has been criticised for not accounting for individual differences in male IPV perpetration and inadequately explaining female IPV perpetration as self-defence (Dutton & Corvo, 2006; Hamel, 2009; Straus, 2008). IPV offenders are a heterogeneous population and gender inclusive approaches consider perpetration to be a result of a complex interplay of biological, psychological, social and contextual factors (Dutton & Corvo, 2006; Straus, 2008). This is supported by research relating to the 'typology' of IPV offenders (Holtzworth-Munroe, 2000; Johnson, 2006). The empirical research has convincingly demonstrated that there is an existence of different types of IPV perpetration involving situational couple violence, controlling behaviours, self-defence behaviours and extreme responses to separation (Johnson 1995, 2008).

'Situational couple violence' was identified as the most common, found amongst community samples and 'intimate terrorism' as the rarer, yet more severe forms of IPV (Johnson, 2008). Gender symmetry occurs across the varying severities of IPV (Ehrensaft et al., 2004; Straus, 2011). Whilst research evidence highlights that IPV is committed equally between genders within Western societies, it is noted this does not transpire across other nations where social roles are deemed less individualised and gender roles unequal (Archer, 2006). Straus (2008) commented that it is crucial to recognise types of IPV perpetrators to inform prevention and policy. Therefore both non-clinical (community) and clinical samples, male and female perpetrators, should be utilised in research.

There is substantial evidence supporting multi-factorial relationships between criminogenic needs associated with IPV (O'Leary, Smith Slep & O'Leary, 2007) and its' developmental nature (Lussier et al., 2009) including that of Dutton's nested ecological model (Dutton, 1995; 2006). This has helped address the limitations in understanding (Bell & Naugle, 2008) and been utilised in treatment programmes in the UK and North America

(Moderate and High Intensity Family Violence Prevention Programme; MI-FVPP & HI-FVPP; Stewart, Gabora, & Hill, 2000 v.1).

Dutton's nested ecological model (1995; 2006) enables the complexities of the aetiology of IPV to be examined and understood within an interaction between four layers. This incorporates; the individual's life factors and biological and psychological characteristics (ontogenic level); the close family unit, interaction between the couple and coping mechanisms in response to IPV perpetration or victimisation (microsystem level; i.e. the intensity of conflict occurring, antecedents and consequences to the abusive incident); the examination of the role of an individual's friends, support network and employment (exosystem level); and lastly, cultural influences and wider societal structures in which a person lives (macrosystem level). The latter can include wider societal beliefs and government laws.

More recently, the General Aggression Model (GAM; Anderson & Bushman, 2002) has been utilised in understanding IPV perpetration (DeWall, Anderson & Bushman, 2011). The GAM is a biological-socio-cognitive model that integrates a range of theories aiming to comprehensively explain the use, and non-use, of violence and aggression. The model incorporates biological and personality factors in addition to the role of arousal, affect and cognition from childhood to present-day (DeWall et al., 2011). This has been utilised in the development of the recently accredited cognitive-behavioural IPV treatment programme in the UK; Building Better Relationships (BBR; see Chapter three).

Risk factors found to be empirically associated with the perpetration of IPV and can be placed within Dutton's (1995; 2006) nested ecological model and the GAM (DeWall et al., 2011) include, but are not limited to; emotional regulation difficulties (Babcock, Jacobson, Gottman & Yerrington, 2000), personality disorder (Varley Thornton, Graham-Kevan & Archer, 2010), lower levels of IQ (Moffitt et al., 2001), peer difficulties (Moffitt et al., 2001;

Theobald & Farrington, 2012), interpersonal dynamics including marital discord, attachment insecurity and jealousy (Allison, Bartholomew, Mayseless, & Dutton, 2008; Medeiros & Straus, 2006b; O'Leary, Smith Slep & O'Leary, 2007; Puente & Cohen, 2003) and alcohol misuse (Gilchrist et al., 2003; Farmer & Callan, 2012). These risk factors are found to be associated with both adult-and adolescent-age IPV offenders (Shorey et al., 2008; see Chapter two).

The role of anger, hostility and aggression within IPV is complex (Norlander & Eckhardt, 2005). A lack of constraint, impulsivity, aggressiveness and negative emotionality in childhood and adolescence have been found to predict adult IPV in robust longitudinal studies (Bersani et al., 1992; Moffitt et al., 2001; Theobald & Farrington, 2012). Whilst some argue that current evidence of anger being solely associated with IPV is limited (Bowen, 2011) others have found it is apparent in some cases of IPV (Eckhardt, Samper & Murphy, 2008). Indeed, not all violent impulses lead on to violent behaviour (Finkel et al., 2009) although a meta-analysis by Stith and colleagues (2004) calculated a moderate effect size between anger/hostility and physical IPV perpetration.

Whilst the above theoretical perspectives highlight the intricacy of factors within IPV, Finkel and colleagues argue that the research has done little to distinguish between risk factors that are violence-impelling, violence-inhibiting or both (Finkel, 2007; Finkel et al., 2009; Finkel et al., 2012). Finkel's 'I^3 Theory' (pronounced "I-cubed theory", 2007) considers the interplay of instigation (provocation), impellance and inhibition. The framework suggests that it may be more common for individuals to experience a violence-impelling force towards an intimate partner following provocation but will only perpetrate physical IPV if these are stronger than violence-inhibiting forces. These forces can be distal, dispositional, relational or situational. This theory therefore aims to understand the interplay between violent impulses and the management of such impulses, enabling a more holistic

approach which incorporates self-regulation literature (Baumeister & Vohs, 2004). Anger is placed as an important dispositional impelling force for IPV perpetration, with self-regulation as an example of a violence-inhibiting force. Such is consistent with research, with high levels of Trait anger and hostility (dispositional aggressiveness) being found among those who perpetrate more severe IPV (Murphy, Taft & Eckhardt, 2007; Norlander & Eckhardt, 2005). Furthermore, a lack of constraint has been identified as a significant longitudinal risk factor for antisocial behaviour, aggression and IPV (Moffitt et al., 2001). The consideration of self-regulation processes has also been found to be an important predictor and moderator in physical IPV perpetration (Finkel et al., 2009) whereby the likelihood of IPV perpetration increases when instigation is strong, dispositional aggression is strong and inhibition (for example, self-regulation) is weak (Finkel et al., 2012). Physical IPV perpetration significantly reduces when just one of these processes trends in the opposite direction. As such, ensuring the accurate measurement of anger experience, expression and control is highly important for understanding IPV, intervention and risk management across the lifespan.

Offence-supportive and aggression-supportive attitudes have been argued to account for more variance of aggressive behaviour than anger alone (Gilbert, Daffern, Talevski, Ogloff, 2013; Gilchrist et al., 2003). Experiences in childhood and adolescents shape an adult's attitudes and perceptions of themselves, others and the world (Bandura, 1986; Prospero, 2006). It has been widely researched but what remains inconclusive is the underlying function of such attitudes, i.e. do these attitudes give permission for IPV or appease feelings of guilt? In addition, are specific attitudes associated with specific types of IPV; physical, sexual, economic, psychological/emotional? Given the dominance of social cognitive theories within the field of aggression, this is an important construct to understand when considering aetiology and intervention (Gilbert et al., 2013).

Victims of IPV are not confined to a specific gender, age- or ethnic group. Therefore, IPV should be viewed from a life-course perspective to enable for further exploration of onset, developmental trajectories, social context and prevalence (Ehrensaft, Moffitt & Caspi., 2004; Moffitt et al., 2001).

Research has consistently found that IPV has far reaching consequences upon adult victims, children, extended families, neighbours and society as a whole. This includes increased risk of mental health problems (Golding, 1999), poorer physical health and poorer cognitive abilities. Such effects were found to have been a consequence of both physical and psychological IPV victimisation across age ranges (Lawrence, Orengo-Aguayo, Langer, & Brock, 2012).

As mentioned above, IPV intervention development should be based upon RNR principles (Andrews & Bonta, 2010a). IPV intervention programmes for both adults and adolescents have been found to yield inconsistent results with regards to effectiveness (Babcock, Green & Robie, 2004; Fellmeth, Heffernan, Nurse, Habibula & Sethi, 2013; Wekerle & Wolfe, 1999; see Chapters two and four). It has been argued that there are other ways to define 'treatment successes' not least with the ever-growing body of research around desistance and questioning 'how' an offender ceases to offend (Laws & Ward, 2011).

The understanding of what constitutes an 'effective' IPV intervention can be achieved through continued rigorous evaluation of interventions. Further, understanding the inter-play of factors associated with IPV at each stage of the life-course can enhance appropriate development of IPV intervention. Such research is imperative in reducing re-offending rates and the risk of harm to victims. It is within this context that this thesis aims to contribute to the on-going research into treatment effectiveness in the field of IPV.

Aim and structure of thesis

The main aim of this thesis is to further academic understanding of intimate partner violence treatment effectiveness and wider practice issues across the lifespan. It aims to pull together research from both adolescent- and adult-aged populations to provide a well-rounded understanding of IPV throughout the life-course and discuss the implications and importance of considering IPV at each life stage. This thesis also aims to further inform academics, practitioners and policy-makers in the consideration of treatment effectiveness and the trajectory of IPV.

Referring to Dutton's nested ecological model, this thesis focusses upon factors found at the ontogenic level (an individual's psychological characteristics) including the role of anger/hostility and attitudes towards violence at adult and adolescent-aged populations. However, discussion will not be limited to these individual characteristics since both individual and relationship (contextual) factors play an important role in understanding IPV (Stith et al., 2004). Consideration will be given to any similarities and/or differences between the two age populations examined.

Chapter two aims to provide an up-to-date review of the literature examining the role of 'offence-supportive' attitudes in the prevalence of adolescent intimate partner violence (AIPV). Furthermore, the goal is to examine such attitudes within specific contexts of AIPV perpetration; physical, sexual or psychological. The discussion draws together the findings of the review, placing them within framework of existing literature and methodological limitations of the studies. The conclusion highlights the continued need for UK based research in this area particularly examining each form of IPV and for primary, secondary and tertiary intervention provision to be available ensuring it is appropriately targeted as highlighted by the RNR model (Andrews et al., 1990).

Chapter three aims to extend current knowledge regarding the effectiveness of IPV intervention programmes in the UK by utilising a sample of adult male offenders based in the community. This retrospective cross-sectional study includes detailed demographic and contextual variables, statistical analyses on both pre- and post-treatment psychometric measures and re-offending and reconviction rates. The author presents these findings as the first evaluation of the Community Domestic Violence Programme ran within Thames Valley Probation Trust.

Chapter four examines and critiques the State-Trait Anger Expression Inventory Second Edition (STAXI-2; Spielberger, 1999) and is used within the authors' research paper as a means of evaluating treatment effectiveness. The rationale for this relates to the consistency of anger/hostility across the lifespan in antisocial behaviour and IPV (Moffitt et al., 2001). Furthermore, the tool is widely utilised among both adolescents and adults and also in assessing treatment outcome with IPV offenders. Conclusions and recommendations are presented with reference made to findings from Chapter three.

The final chapter, Chapter five, draws the issues raised throughout the thesis together in relation to theoretical implications, IPV prevention, intervention programme development and risk management across the lifespan. This includes reflection on the complex interplay of factors involved in IPV across adolescent and adult perpetrators; the need for multiagency co-operation to include the immediate family unit in these cases and the importance of adopting a more targeted approach for offenders engaged in IPV intervention programmes. Consideration is also given to the limitations within this thesis and implications for future directions in research and practice are discussed.

Chapter Two

The association between adolescents' attitudes towards adolescent intimate partner violence and reported perpetration and/or victimisation among community samples: A Systematic Review.

Rationale for Chapter Two

Chapter one introduced evidence of the increasing prevalence of AIPV (Barter et al., 2009), intergenerational links between witnessing IPV in the home and subsequent IPV perpetration (Lussier et al., 2009; Narayan et al., 2014; Roberts et al., 2010; Stith et al., 2000) and highlighted IPV as being a life-long issue for some (Moffitt et al., 2001). Given the evidence for IPV across the lifespan, a fuller understanding of such experiences in adolescent years may be informative in developing more effective treatment programmes targeted at adult IPV offenders in addition to adolescent targeted prevention and intervention. An increasing amount of research investigating aetiology, mediating factors and prevalence has been carried out on college-aged samples relating to risk factors (Shorey et al., 2008). However, less has been conducted on younger high-school aged populations.

The role of cognition, including attitudes towards violence, has been found to be an important mediator in aggressive and violent behaviour (Anderson & Bushman, 2002). Indeed there has been empirical support indicating that attitudes towards violence are directly associated with, and may predict, AIPV (Foshee et al., 2001). However, inconsistent findings and questions relating to the complex relationship between attitudes and behaviour remain Bowen, Gilchrist & Beech, 2008; Leen et al., 2013).

Abstract

This review aims to systematically analyse existing literature that has explored the relationship between attitudes towards dating violence and reported perpetration and/or victimisation of AIPV. Electronic databases were searched using a systematic approach. Pre-defined inclusion and exclusion criteria were applied to identify relevant literature. Studies were then quality assessed according to a pre-defined protocol. The data was extracted from the included studies and analysed descriptively.

In total, seven studies were included in this review. Five out of the seven studies demonstrated modest but significant relationships between attitudes and dating violence perpetration and victimisation. Two studies partially supported the association. Prevalence rates were identified between 15% - 39% and 26% - 68% for males and females respectively. The association between attitudes supporting AIPV and abusive behaviours was more prevalent among males. Younger aged high-school students and 'high-risk' community populations were more accepting of AIPV with the latter presenting with higher rates of perpetration and victimisation. The cross sectional nature of each study does not allow for understanding of the direction or results or causality.

The results show some consistency that the acceptance and justification of violence is related to the use of dating violence, however methodological differences between studies and limitations within studies mean it is not possible to draw unequivocal conclusions pertaining to each category of AIPV; physical, sexual, psychological/emotional. Implications for research, assessment and intervention are discussed.

Introduction

Research into the developmental risk factors associated with adult intimate partner violence indicates that the phenomenon is not confined to adult relationships. It has been found in a number of studies that AIPV is increasingly common (Jackson, 1999; see Chapter one). Most of the empirical evidence on AIPV is derived from North America although more studies are being conducted in the UK (Barter et al., 2009). As is the case with adult IPV, the consequences of experiencing AIPV are far reaching with victims reporting increased levels of anxiety, depression, physical health complaints (Haynie et al., 2013), distancing from friends and increased risk for drug and alcohol misuse (Foshee, McNaughton Reyes, Gottfredson, Chang & Ennett, 2013). As young adolescents enter into adulthood, what is learned at this stage may be the foundation of future longer-term behaviour (Wekerle & Woolfe, 1999).

Historically, there have been difficulties associated with the study of AIPV (Note: the North American terminology for AIPV is 'dating violence' (Barter et al., 2009) but AIPV is used throughout to maintain consistency within this thesis). A main obstacle includes a lack of a standardised operational definition as to what constitutes violence within an adolescent intimate relationship. According to the Centre for Disease Control (CDC), AIPV is defined as: "the physical, sexual, or psychological/emotional violence within a dating relationship" (CDC, 2008). In an extensive review of the literature across Europe and North America, Leen et al. (2013) found psychological/emotional AIPV to be the most and sexual violence as the least prevalent. Some researchers have proposed that the use of sexual violence within relationships is inherently different to the use of psychological and physical abuse (Sears, Byers & Price, 2007). This is evident in the definition proposed by Glass et al. (2003) who defined AIPV as the "perpetration or threat of an act of violence by at least one member of an unmarried couple within the context of dating or courtship (same sex or opposite sex)" (p.

228).This definition presents as reductionist and inconsistent with definitions of adult IPV. It is imperative to explore each of these forms directly to enhance understanding of the phenomenon at specific age stages and across the life-span. This is not without challenges given the increased sensitivity associated with participants under aged 18 years. Under-reporting also has an impact on obtaining accurate prevalence rates and honest accounts of attitudes and behaviours (Jackson, 1999).

Research findings from studies investigating risk factors associated with AIPV have been largely consistent with those associated in IPV perpetration (Shorey et al., 2008; Shorey et al., 2012). Those that have been identified as associative with or predictive of AIPV include; peer influences (Barter et al., 2009; Foshee, Linder, MacDougal, & Bangdiwala, 2001; Leen et al., 2013; O'Keefe, 2005; Vagi et al., 2013), angry temperament and poor anger control (Baker & Stith, 2008; Harper, Austin, Cercone, & Arias, 2005). Williams, Connolly, Pepler, Craig and Laporte (2008) found that peer delinquency, peer aggression and high acceptance of dating aggression significantly predicted adolescents' recurrent aggression in new dating relationships. When examining the literature regarding risk factors, it was evident that the studies used can vary between high school-aged and college-aged samples (Shorey et al., 2008). Caution is required when interpreting these findings in accordance to AIPV as findings from college-aged samples may not be generalisable to younger high school-aged samples.

There have been some gender differences identified amongst risk factors associated with AIPV (Baker & Stith, 2008; Leen et al., 2013; Prospero, 2006) including alcohol misuse being a predictor of AIPV amongst female perpetrators but not male perpetrators (Foshee et al., 2001). Gender differences should not be overstated given the similarities across both sexes regarding risk factors for antisocial behaviour more generally, despite males engaging in more of it (Moffitt et al., 2001). There is little evidence that males are more vulnerable to

any set of risk factors than females and indeed, IPV perpetration has been found to be of similar levels across sexes (Moffitt et al., 2001). Furthermore, in cases of life-course persistent offenders of anti-social behaviour, males and females are found to share similar risk factors including family adversity, cognitive deficit, peer rejection, hyperactivity, weak temperament and poor discipline (Moffitt et al., 2001). It is noted that female life-course persistent offenders are very rare. The risk factors examined within this chapter highlight male trends due to the focus on male perpetrated IPV within this thesis.

Evidence pertaining to attitudes condoning general violence and the perpetration of AIPV has found that higher levels of violence are inflicted upon victims if pro-violent attitudes are held (Munoz-Rivas, O'Leary & Gonzales, 2007; Williams et al., 2008). Further, attitudes that are accepting of violence appear to be more prevalent for boys opposed to girls (Foshee et al., 2001). A comprehensive international review authored by Leen et al. (2013) examined prevalence of AIPV, dynamic risk factors and the efficacy of AIPV primary interventions. Within this review, six longitudinal studies were included that investigated the causal relationship between attitudes towards violence upon AIPV perpetration. The review summarised that the presence of pro-violent attitudes was not a stable long-term predictor of AIPV since the results suggested more of a concurrent association. The authors suggest that this may be understood situationally and that attitudes may fluctuate to reflect behaviour in order to reduce cognitive dissonance. Indeed, adolescents may often not understand the differences between abusive and healthy dating behaviours, i.e. perceiving jealousy as a positive sign of love (Callahan, Tolman & Saunders, 2003). Further, gender differences were evident including that acceptance of dating violence was only a longitudinal predictor for boys in certain studies and girls' use of psychological and physical aggression was found to be much higher (Leen et al., 2013).

Bethke and DeJoy (1993) found that only half of adolescents would terminate a relationship following a violent act. This could be considered in the context of findings in the Dunedin study that antisocial males and females are more likely to select mates who are disadvantaged and hold anti-social values (Moffitt et al., 2001). It is suggested in the literature that there is a specific deficiency of studies investigating the role of perception and attitudes among younger adolescents concerning dating violence (Prospero, 2006).

The issues highlighted thus far refer to dynamic risk factors (i.e those which can change in either direction; Andrews & Bonta, 2010b). With regards to static risk factors (historical factors that are unchangeable; Andrews & Bonta, 2010b), Chapter one highlighted the predictive nature of witnessing violence within the family home (Theobald & Farrington, 2012). Predominantly this has been examined within the context of adult IPV and must be considered as part of a multifaceted developmental model of antisocial behaviour that is influenced by several factors within an individuals' environment in addition to biological and personality factors (Lussier et al., 2009). More recent studies have explored the significance of witnessing violence within the family home as a risk factor among adolescent samples and has been found to predict and mediate AIPV (Narayan et al., 2014; Temple et al., 2013). It is noted that not all studies have found significant associations (Schwartz, O'Leary & Kendizora, 1997). Whilst static risk factors cannot be changed through intervention as dynamic risk factors can (Andrews & Bonta, 2010a), the presence of this risk factor is important when considering the development of attitudes that condone general violence and violence towards intimate partners in relation to social cognitive theories (Bandura, 1986).

Research has supported the concept of a developmental trajectory of the perpetration of intimate partner violence in adolescence to adulthood although many studies have mainly focussed upon college-aged students and there is a gap in the literature about the aetiology

and epidemiology specifically relevant to a high school-aged sample (Wekerle & Wolfe, 1999).

Social cognitive theories of AIPV

A range of social-psychological theories have been utilised to explain interpersonal violence and the role of perception and attitudes. Social Learning Theory (Bandura, 1986) has been applied to IPV, and also AIPV. When applied to AIPV, social learning theory suggests that AIPV develops as a result of adolescents learning inappropriate dating norms through the observation of people (and role models) in their lives. This theoretical element is applicable to research findings related to the developmental pathway of risk of perpetrating AIPV and subsequent IPV in adulthood as highlighted in Chapter one, including witnessing IPV in childhood (Narayan et al., 2014; O'Keefe, 2005) and peer influences (O'Keefe, 2005).

According to social learning theory, an adolescent's continuation of abusive behaviour is dependent upon positive and negative reinforcements and punishments that follow the behaviour. Further, the perception of this received consequence leads to a continuation of that behaviour. In the context of violence towards intimate partners, attitudes condoning or justifying the use of aggression may therefore be further influenced by the perceived consequences to the intimate partner relationship. Indeed, there is some criticism that such social-cognitive theories do not take into account individual factors such as genetic influences and psychopathological factors (Ferguson & Dyck, 2012).

Theories examining individuals' perceptions of interpersonal relationships and motives have been widely researched. Hostile Attribution Bias refers to a tendency to perceive hostile intent in the actions of others, regardless of its true presence (Dodge, Prince & Newman, 1990). This theory links to AIPV as it could serve to justify the use of aggression in response to a perceived threat. Prospero (2006) highlighted the importance in

understanding the role of perception of violence among adolescents and its development. In particular the important role that negative peer behaviour may play in adolescent's development of attitudes towards dating violence was emphasised within this population. This study of 89 middle-school aged students (M = 13.46 years) utilised 'common dating scenarios' that were devised for the purpose of the research. Participants were asked to comment on the protagonists' perception of the scenario and then the protagonists' behavioural response which included aggressive and non-aggressive options. Prospero found a moderate correlation (r = .338) between perception and behavioural responses within the dating scenarios. However, the presence of several other factors is inherent given the level of association. Males and females did not significantly differ in their perceptions although males reported expecting a higher level of aggressive behavioural responses. The limitations of this study is that the middle school-aged sample consisted of individuals who were identified as "at risk" and thus enrolled in a leadership programme incorporating violence prevention. The demographics and characteristics of these individuals may be different from that of other middle school-aged individuals who were not identified as "at risk". Given that there was no control or comparison group, it is difficult to make sound conclusions regarding the ability of the study to be generalised further than its participants. Questions remain as to 'why' these situations were associated with aggressive perceptions and the underlying mechanisms that can lead to justifying aggressive behaviours including the role of attachment, angry temperament and jealousy (Follingstad, Bradley, Helff & Laughlin, 2002). Further understanding of what triggers an individual's pro-violent attitudes is required to further understand the development and perpetration of AIPV, and adult IPV.

Cognitive Dissonance Theory (Festinger, 1957) refers to the relationship between cognitions and behaviour. The theory proposes that an individual is motivated to have consistency between their cognitions and behaviour. Where inconsistency occurs between

cognitions, attitudes or behaviours (dissonance), the individual feels discomfort and may adapt, modify or change one of these areas to ensure consistency. This can lead to maladaptive and/or unhelpful behaviours. Schumacher and Smith Slep (2004) applied this theory to examine the inconsistency between self-reported attitudes and behaviours within dating relationships and whether this dissonance predicted change in abusive behaviours. This study of 398 high school students (aged between 15 and 18 years) identified 58% of males as 'dissonant' with regard to verbal aggression and 47% as dissonant with regard to jealous tactics. The sample was compared from Time 1 to Time 2 (three months apart). Verbally aggressive behaviours at Time 1 predicted the perpetration of the same behaviours at Time 2 (accounting for 2% - 4% of the variance) and cognitive dissonance status enhanced the prediction, accounting for up to 6% of the variance. Further, the study found that attitudes that were non-supportive of psychologically aggressive behaviours predicted a decrease in the engagement in such behaviours. This demonstrated preliminary evidence that dissonance may be a catalyst for change.

Researching AIPV is highly complex, given the multifaceted nature of perpetration (Shorey et al., 2008). Whilst the theoretical background provides potential understanding of the problem, the application of such theories within AIPV research is limited at present. Scholars have placed significance upon attitudes as a target for change in AIPV prevention programs, with it often being evaluated as the sole outcome measure. There is some evidence to indicate that treatment programmes (e.g. The Safe Dates Program; Foshee, 1996) have produced promising results in reducing cognitive distortions relating to dating violence in the short term (one month follow up). However, these changes did not continue past the one year mark. The amount of research currently conducted in developing prevention programmes and evaluating their long-term effectiveness is far behind research relating to AIPV risk factors (Shorey et al., 2012). Whilst such evidence is hesitantly positive, there remains a debate as to

the interaction and predictive stability of pro-violent attitudes being indicative of violence behaviours (Leen et al., 2013; Shorey et al., 2012; see Chapter four). It is therefore essential for further research to explore the specificity, significance and function of attitudes among adolescents towards AIPV, and indeed adults in relation to IPV, to assist in its prevention and the protection of adolescents.

Aims

This review aims to systematically explore and analyse existing research that has investigated the relationship between attitudes among community samples of adolescents towards IPV and the prevalence and nature of violence within adolescent dating relationships. The review aims to enhance understanding of the underlying situational mechanisms of attitudes within specific types of AIPV as influenced by the work of Leen et al. (2013).

Leen et al's (2013) review is one that examined prevalence of each specific type of AIPV; physical, sexual and psychological and noted a dominance of research into physical AIPV with fewer studies into other forms of AIPV with many suffering from methodological differences. Additionally, when examining risk factors for AIPV, attitudes towards violence were found to be more strongly linked at the concurrent, situational, level. When referring to Finkel's I Theory (Finkel 2007; Finkel et al., 2012; see Chapter one), the processes that occur at the situational level of an individual are arguably inherent to engaging in, or avoiding, aggressive and/or abusive behaviour. Therefore, this review will examine attitudes and its' association to all three AIPV types which Leen et al. (2013) did not specifically examine.

All studies included in the review explored the relationship between attitudes towards AIPV as a primary focus or as a part of a multifaceted analysis. The outcome is defined as

adolescents' reports of perpetration and victimisation of AIPV. More specifically, the objectives of this review were to establish:

1. The extent of association of pro-violent, acceptance and/or justification attitudes with perpetration and/or victimisation of each type of AIPV; physical, sexual and psychological/emotional.
2. Whether attitudes vary depending on the specific categories of AIPV perpetration and/or victimisation, including physical, sexual and psychological/emotional abuse.

Method

Scoping search strategy

An initial scoping search was carried out to identify any existing literature reviews on the role of attitudes among adolescents towards dating violence. In order to conduct the scoping search, terms were entered into the Cochrane Database of Systematic Reviews (CDSR) and DARE databases to identify systematic reviews or meta-analyses relating to this topic (All years: 3rd May 2014; see Appendix one for details of search terms used). One review of note was found in that it was considered to be informative to the aims of the current review (Fellmeth et al., 2013).

Fellmeth et al. (2013) conducted a systematic review and meta-analysis examining the effectiveness of AIPV interventions which included exploring attitudes towards AIPV as a secondary outcome. The meta-analysis found a standardised mean difference (SMD) of 0.06 (95% CI -0.03 - 0.15) using a fixed effects model. Due to heterogeneity, a random effects model was also conducted and yielded an SMD of 0.08 (95% CI -0.06 - 0.22). The authors state that whilst this suggested slightly improved attitudes in relation to AIPV, the CI level included the possibility of the interventions having no effect or indeed worsening

participants' attitudes. Two further reviews with regard to the effectiveness of AIPV prevention programmes (Foshee et al., 2005, Wekerle & Wolfe, 1999) highlighted small but significant results on a reduction pro-dating violence attitudes following completion of dating violence prevention programs. This is considered consistent with results pertaining to the effectiveness of IPV programmes targeted at male adults (Babcock et al., 2004; see Chapter three).

PsycINFO was also searched using terms that enabled a restriction of results to include only literature reviews and/or meta-analyses (see Appendix one for details of search terms and syntax). The search of PsycINFO yielded 79 results which were filtered by hand.

Main literature search strategy

Following the return of results in the scoping search, a standardised search was carried out in electronic bibliographic databases to identify relevant studies for the current systematic review: EMBASE (1988-2014, Week 20); MEDLINE (1946 to 2014, May Week 1); PsycINFO (1987-2014, May Week 2). The search of electronic databases was conducted on 17th May 2014. Search results were limited to articles published within the past 20 years and restricted to the English language. A summary of the search terms and syntax can be found in Appendix two.

Initial search results were saved to RefWorks then screened by hand to identify the relevance of the studies. This process consisted of viewing the title and abstract of each article and removing them if they were found to be irrelevant to the current review. If there was insufficient information gained from viewing the abstract about the eligibility of the study, it was reviewed by accessing the full text. Duplicates were also removed. In addition to electronic database searches, the reference lists of relevant studies were also examined with

the aim to identify relevant research papers as outlined by the current review's objectives. This process identified three further studies that had not appeared in the initial search. One international researcher was successfully contacted to request full access to an identified article that was unavailable to the author electronically. Ideally, an examination of all identified hard copied journals would have been conducted to increase the depth and breadth of the search although this was not possible for the current review.

Study Selection

The following criteria were used to assess study eligibility for the present systematic review (see Appendix three). This criterion was applied to all remaining studies and was determined by one researcher.

Population:	Young to Late Adolescents (male and female or solely male) aged between 11 and 21 years of age Population from US or UK Community samples
Intervention:	Exposure to assessment of attitudes towards dating violence
Outcomes:	Reported dating violence incidents (perpetration only or perpetration and victimisation)
Study Type:	Cross Sectional/ Cohort, Quantitative / Mixed design
Exclusion:	Editorials, dissertations, commentaries; juvenile offender/forensic/psychiatric populations or mixed samples; studies that focus upon AIPV as part of wider aggressive behaviours; studies that focus on an exclusively male or female population; longitudinal design; intervention effects; and studies where the outcome measure was vignettes

as opposed to self-reported AIPV.

Language: English only

Initially the inclusion criteria consisted of a population between 11 and 18 years old, however some studies included a minority of individuals older than this. As a result, the age was increased to include a maximum age of 21 however the mean age within all studies remained below the age of 18 years. Implications of this age bracket will be discussed. No explicit comparator was identified due to exclusion of treatment outcome studies in line with the aim of the current review.

Longitudinal studies were excluded from this current review following from Leen et al.'s (2013) findings that attitudes condoning violence were not identified as a stable predictor. Consequently, the aim of this review was to examine attitudes from a concurrent perspective focussing upon disposition of an individual to hold such attitudes and the situations whereby they are associated with AIPV perpetration and/or victimisation. Furthermore, qualitative studies were excluded to reflect the intention of synthesising quantitative results as a preliminary investigation into this specific area. Time constraints meant not being able to use Thematic Analysis or Interpretative Phenomenological Analysis to synthesise results from qualitative data however this is for consideration in future reviews.

The outcome measure was identified as self-reported real-life perpetration and/or victimisation of AIPV. It is of note that during the search stage, some studies included an outcome measure of participant's response to dating violence vignettes. Whilst these studies utilised detailed measures of attitudes and perception (for example, Prospero, 2006), they did not contain outcome measures identifying self-reported perpetration and/or victimisation and were therefore excluded.

Data extraction

The data extraction process was conducted by one reviewer using a standardised data extraction form applied to each study (see Appendix four). This enabled various information from the studies to be recorded including; re-verification of study eligibility (target population and inclusion/ exclusion criteria), study design, aims, sampling, assessment measures, attrition rates, statistical analysis and results. Clarification of any unclear issues through contact with authors of studies could not be conducted due to time constraints and may have therefore impacted upon the conclusions.

The initial search of the three electronic databases obtained 1128 results. Of those, 281 were duplicates and therefore removed with a further 606 excluded due to being identified as not relevant to this review. Subsequently, 244 studies were measured against the inclusion criteria. A total of 237 studies did not meet the criteria resulting in seven studies being taken forward for quality assessment. No studies were excluded on the basis of poor quality as is described below (see Figure 1 for a flowchart detailing the search results). A summary of the characteristics of the studies are shown in Table 2.1. Where studies examined a variety of factors, only hypotheses, measurements and results relevant to the current review's aims were extracted. One study (Reeves & Orpinas, 2012) utilised a mixed design. Only the data from the quantitative element of this study was extracted.

Assessment of methodological quality

The seven studies found to meet the inclusion criteria were subject to a methodological quality assessment. The checklist was adapted from The Critical Appraisal Skills Programme (CASP; see Appendix five). The studies were then classified for total methodological quality on the basis of: design, selection bias, measurement bias, attrition bias and outcome/result bias. Each item was scored on a three point scale: Item fully met = 2; Item partially met = 1; Item not met = 0. There was one item that was reverse scored. Items in the quality assessment

checklist that were not available or undefined in the study were not scored and identified as "unclear". The maximum score was 58.

Six out the seven of the studies included in this review are of non-experimental cross-sectional design (one is a cross-sectional cohort study; Simon, Miller, Gorman-Smith, Orpinas & Sullivan, 2010). As participants were observed at one point in time, the studies cannot infer a causal affect. This weakens the quality of the study design. No control groups were used in the studies, which further reduce the reliability of the findings. The cut off percentage was 60% as this was considered reasonable to ensure only good quality studies were included within the parameters of limitations above. A high percentage score in the quality assessment indicates higher reliability and validity however this was considered within the limitations of the aforementioned weaker design. All seven studies scored above the cut off, ranging from 64% - 78% and therefore included in the review. Table 2.2 provides a summary of the quality assessment of each study.

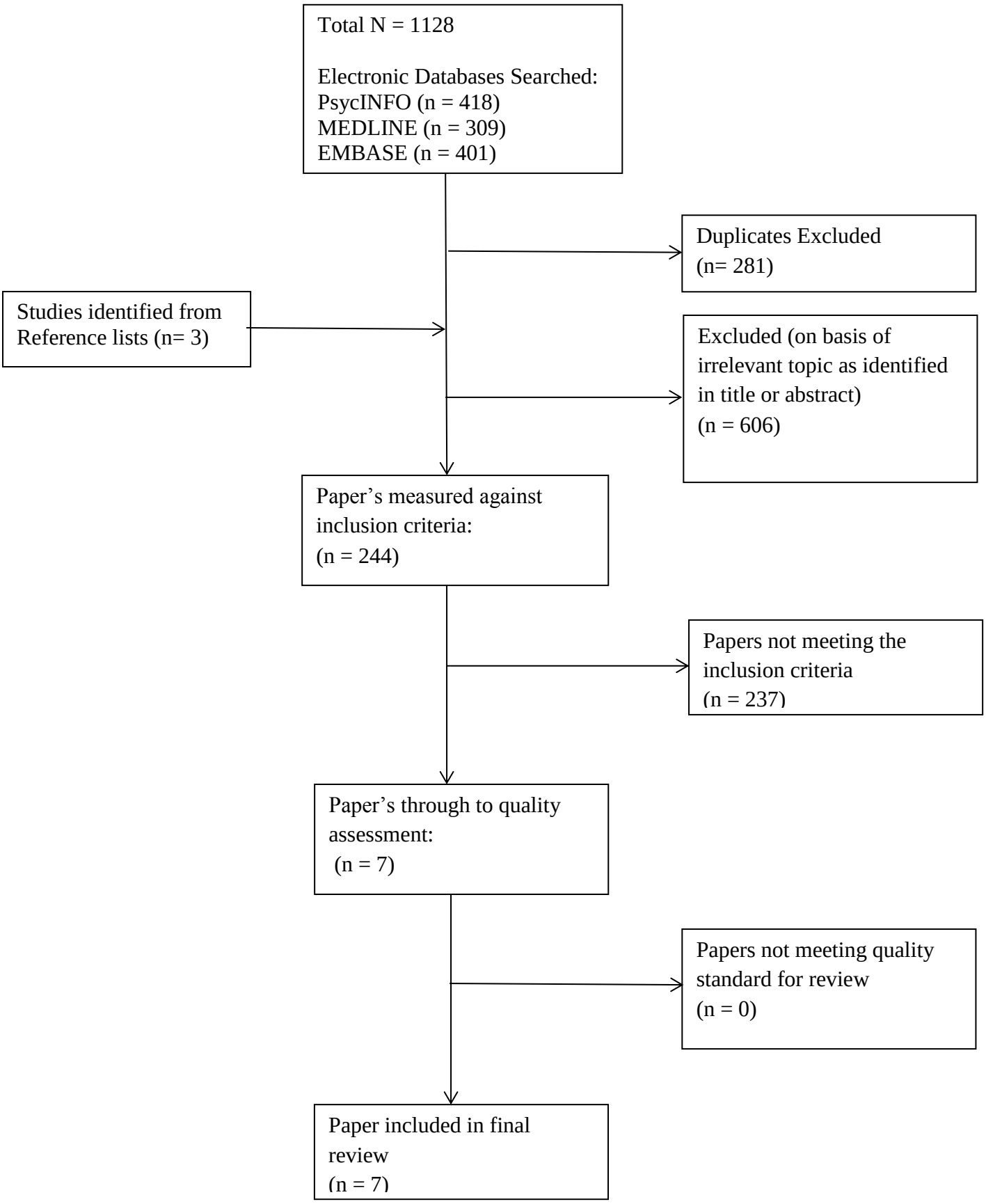

Figure 1: *Flow diagram of the search process*

Table 2.1

Studies exploring attitudes justifying AIPV and AIPV perpetration and victimisation.

Authors, (Year), Country	Type of Study	Study Aims	Participants	Exposure Assessments Used	Outcome Measures	Findings
Ali, Swahn, and Hamburger (2011) North America	Cross-Sectional	To examine the associations between attitudes that support physical dating violence against boys hitting girls and vice versa and experiences with **physical dating violence** perpetration and victimization.	*n*=2888 in an urban school district. Identified as 'high-risk community'. Male: 1383 (47.9%) Female: 1505 (52.1%) Age: 12-18 years Grades 7 (28.6%), Grade 9 (28.6%) and Grades 11/ 12 (42.8%) Ethnicity: Hispanic (44.5%); African American (27.4%); Caucasian (23.6%); Other (4.5%).	Dating violence attitudes assessed using a 10-item scale modified from Foshee et al. (2001). 5 items included **attitudes of males' physical violence towards females and 5 examined females' physical violence towards males** using 4 point Likert scale.	Modified measure (Foshee et al., 1996) for **physical perpetration and victimisatio**n in dating relationships and peer relationships. Response options included never, *1–3 times*, *4–9 times*, and *10 or more times* although this was dichotomised.	1. Results from the multivariate logistic regression analyses show that attitudes supporting boys hitting girls (Adj. *OR* = 1.44; 95% CI: 1.12–1.84) as well as attitudes supporting girls hitting boys (Adj. *OR* = 1.47; 95% CI: 1.14–1.88) were statistically associated with **physical dating violence** perpetration after controlling for potential confounders. 2. Gender specific attitudes associations: Attitudes condoning boys-to-girls violence was associated only with boys perpetrating physical AIPV and victimisation of boys. Attitudes condoning girls to boys was only associated with girls perpetrating physical AIPV. 3. Additionally, holding attitudes that support girls hitting boys was significantly associated with physical dating violence victimization for both boys and girls.
Chase, Treboux, O'Leary and Strassberg (1998) North America	Cross-Sectional	To examine the specificity of the use of **physical AIPV** and its justification in response to interpersonal	*N*=95 Adolescents attending community (school) drop-out prevention programme for problem behaviour.	Thoughts about Relationships Questionnaire (TARQ: developed for the study). Included 12 hypothetical vignettes (three situations presented for four types of	Adapted version of CTS (Straus, 1979) Reports of **perpetration and victimisation of physical dating**	1. 33% males and 68% of females reported using **physical violence** against their current partner. 38% males and 79% of females reported using **physical violence** against ex-partner. 2. Overall, justification for use of aggression

Authors, (Year), Country	Type of Study	Study Aims	Participants	Exposure Assessments Used	Outcome Measures	Findings
		problems. Examined whether it is specific to current partner, general to dating relationships, or a general age-mate aggression issue (same-sex peers and opposite-sex partners).	Identified as high risk. Male: n=61(64%) Female: n=34 (36%) Age range: 14-19yrs. Male age: *M* = 17.32; *SD* = 1.31. Female age: *M* = 16.78, *SD* = 1.31. Ethnicity: 65% Caucasian; 14% African American 9% Hispanic; 12% "Other".	age-mates) Respondents are to rate **extent to which use of physical aggression was justified**. 4-point Likert scale.	**violence** within a current intimate partnership, previous partners and same sex peers.	against an age-mate was associated with use of aggression. Males **physical AIPV** against current partner was significantly associated with their justification of aggression against a previous partner (r_{pb} = .35, p = < .01). 3. Moderate correlation between justification for and use of aggression against the CDP (r_{pb} = .28, p = < .05). 4. **Physical AIPV** was found to be associated across all dating relationships (current partner and ex-partner) for males (φ = .73, p < .001) and partner-specific for females.5. Justification of physical AIPV against a previous partner was predictive of physical AIPV towards current partner (p = < .01) but justification of current partner physical AIPV did not predict engaging in physical AIPV against that partner. Combination of the two justifications above predicted current partner physical AIPV, $F(2, 58) = 3.85$, p = < .05, R = .34.
Feiring, Beblinger, Hoch-Espada and Haworth (2002) North America	Cross-Sectional	To examine use of **physical and emotional AIPV** and attitudes condoning such behaviours and role of	*N* = 254 high school students. Grades 9 – 10 (*n*=89) and 11-12: (*n*= 165). Female: *n*=160 (63%)	Relationship Attitudes Survey for Adolescents (RASA; Deblinger et al., 2000) **measuring attitudes concerning sexual interactions and intimate relationships**	Conflict in Relationships measure (CIR; Wolfe et al., 1998) was used to index **physical and emotional abuse.**	1. Aggressive behaviours and endorsement attitudes/ dysfunctional sexual attitudes (DSA) were not common in dating relationships. 1/4 reported perpetrating at least one of milder forms of aggression, 1/5 victimized. Over half reported **emotional abuse victimisation.** Girls reported higher

Authors, (Year), Country	Type of Study	Study Aims	Participants	Exposure Assessments Used	Outcome Measures	Findings
		attachment and emotional styles in relation to these behaviours and attitudes.	Male: *n*=94 (37%) Age Range: not clearly specified however included students in Grade 9-12 (14yrs -18yrs traditionally). Ethnicity: 79% European; 17% African-American; with remaining consisting of Hispanic, Asian and 'other' ethnicity categories.	**(such as tolerance of aggression and healthy relationship attitudes).** Test of Self Conscious Affect for adolescents (TOSCA-A; Tangney et al., 1991). Vignettes examining proneness to guilt/shame and externalisation of responsibility in situations relating to peers. Behavioural Systems Questionnaire exploring beliefs about friends/partners and attachment style.	**Perpetration and victimisation of physically abusive behaviour** in last year **Victimisation of emotional abuse in past year.**	levels of **physical AIPV** (χ^2 (1, N = 254) = 6.1, $p \leq .01$).Those in Grade 11-12 more likely to report victimisation of emotional abuse. 2. Boys more likely to hold attitudes that endorse violence (F(1, 253) = 43.4, $p \leq .0001$) and hold dysfunctional sexual attitudes compared to girls (F(1, 253) = 15.3, $p \leq .0001$). Those in Grades 9-10 held higher dysfunctional sexual attitudes and boys that age held lower healthy relationship attitudes than Grades 11-12. 3. No strong relationship between behaviours and attitudes, although two were significant: perpetration of **physical violence** was positively related to DSA (r = .24, p < .01) and negatively related to healthy attitudes (r = -.16, p < .01). Correlation for the former was stronger for boys (r = .43, $p \leq .01$). 4. The first canonical function for the relation between emotional styles and relationship attitudes was significant for boys (R_c = .47, χ^2 (9, N = 160) = 34.0, $p \leq .0001$) accounting for 22% of variance. For males, externalizing responsibility for harm was related to higher levels of DSA. More guilt and less shame were linked with healthier

Authors, (Year), Country	Type of Study	Study Aims	Participants	Exposure Assessments Used	Outcome Measures	Findings
						relationship attitudes.
O'Keefe (1997) North America	Cross-Sectional	To examine factors that best predicted infliction of **physical dating violence** using a social learning theory based model.	1,102 high school students approached. Final sample: *N*=939 Male: 385 (41%) Female: 554 (59%) Age Range: 14-20 yrs (*M* = 16.9 yrs). Ethnicity:53% Latino, 20% White, 13% African American, 6.7% Asian American, 7% "other".	Demographic Questionnaires **A justification of Violence Scale** adapted from Margolin and Foo (1992). A range of situational variables and contextual variable data was obtained. Measures included a questionnaire examining witnessing of inter-parental violence; history of aggression; violence at school or in the community and various measures examining relationship characteristics (seriousness, number of partners, length of time dating).	Modified version of Conflict Tactics Scale (CTS; Straus, 1979) included reports of **perpetration and victimisation of physical dating violence within an intimate partnership. This included a sexual aggressiveness item.** Participants asked to complete main reasons they perpetrated AIPV using Follingstad et al.'s (1991) list of 13 possible motives for dating violence including anger, jealousy, and need for control.	1. Prevalence of **physical AIPV** perpetration: 43% female versus 39% male. 2. Attitudes justifying **male -to-female physical AIPV** found to be correlated to **physical AIPV infliction by males** (r = .42, p < .05). Also found to be a significant predictor in regression equation at p < .05. 3. Significant gender differences in justification of violence attitudes: **Males physical AIPV** were more likely to inflict dating violence against a partner when they believed male-to-female violence was justifiable. 4. Both sexes more accepting of female-to-male violence (Girls; M = 13.6, SD = 4.9, Boys; M = 13.5, SD = 4.7) 5. Significant gender differences in motives for Anger (χ^2 = 10.2, p = .001; females more likely), to get control (χ^2 = 10.1, p=.001; males more likely) and self-defence (χ^2 = 9.9, p=.002; females more likely). Jealousy motive third most frequent for both sexes.
Reeves and Orpinas	Cross-Sectional	To examine the role of social	*n*=624 male and female Grade 9	Dating norms measure adapted from Foshee et al.	Modified measure adapted from Foshee	1. More participants reported support for girls hitting boyfriends than boys hitting

Authors, (Year), Country	Type of Study	Study Aims	Participants	Exposure Assessments Used	Outcome Measures	Findings
(2012) North America		norms and their association with **physical aggression in dating relationships**, specifically the norms that support male-to-female and female-to-male dating violence.	students (aged between 14 and 15 years). Ethnicity: 47% White: 38% African American; 11% Latino; 5% Other.	(2005); Four items measuring **attitudes regarding male-female physical aggression and four items measuring female-male physical aggression.** 4 point Likert scale: 1 = strongly disagree – 4 = strongly agree.	et al. (1996) **examining perpetration and victimisation of physical aggression within dating relationships.** Asked to report prevalence relating to **seven types of physically aggressive behaviour** within last three months.	girlfriends (one third versus one sixth). Overall acceptance of aggression was low. 2. Gender differences and dating status in the level of **physical dating violence norms** was not significant after controlling for race/ethnicity and dating status. 3. African American students support for violence from girls to boyfriends significantly higher. Support for violence from boys to girlfriends was significantly higher for all non-White students. 4. Association between attitudes and behaviour was weak (and non-existent in some cases) for girls and strong for boys.
Sears, Byers and Price (2007) CANADA	Cross-Sectional	To examine the co-occurrence of and risk factors (including attitudes) for **reported use of psychologically, physically, sexually abusive behaviours** within dating relationships.	*N* = 633from Four Canadian high schools in Grades 7 (*n*=192), 9 (*n*=193) and 11 (*n*=248). Male: *n*=324 (51%). Female: n=309 (49%). Age Range: 12 to 18 years (*M* = 14.64 yrs). Race/Ethnicity:	Demographic information. **Attitudes Towards Women** Scale for Adolescents (AWSA; Galambos, Petersen, Richards & Gitelson, 1985). Scored on a 4-point scale with higher scores indicating less accepting attitudes towards traditional roles. Six scales assessed attitudes towards dating	Modified Version of CTS (Straus, 1979) to assess adolescent's **use and experiences of physical and psychological/ emotionalAIPV** within dating relationships. **Sexual AIPV perpetration and victimisation:** Modified version of the Sexual	1. Both genders liberal in attitudes towards women's roles and low acceptance of any type of violence within dating relationships. 2. Prevalence: Male abusiveness; 43% overall; 35% **psychological,** 15% **physical** and 17% **sexual.** Female abusiveness: 51% overall; 47% **psychological**, 28% **physical,** 5% **sexual.** 3. 43% of boys and 51% of girls reported having used at least one form of abusive behaviour toward a dating partner; and that 19% of boys and 26% of girls had used two or more forms of dating violence.

Authors, (Year), Country	Type of Study	Study Aims	Participants	Exposure Assessments Used	Outcome Measures	Findings
			classified as "vast majority were Caucasian". 85% English Canadian; 6% French Canadian; 6% Native Canadian.	violence (Price, Byers & The Dating Violence Research Team, 1999); Separate **scales assessed attitudes towards male and female psychological, physical and sexual dating violence.** 5-point scale with higher scores indicating greater acceptance of abusive behaviour. Measures examining fear of family violence and peer use of violence used.	Experiences Survey-Revised (SES-R; Koss, Gidycz & Wisniewski, 1987). Respondents used a 3-point scale to identify if they had engaged in the behaviour; never (0); once (1); more than once (2) or been the victim of the behaviours.	4. **Most common type of AIPV behaviour used was psychological for males and both psychological and physical for females.** More girls than boys reported using more than one type of abuse, χ^2 (2, N = 633) = 6.08, p = < .05. 5. **Boys: more likely to perpetrate all forms of AIPV when** they; were victims of all forms of AIPV, held more accepting attitudes of all forms of male use of AIPV and more traditional values of women's roles. Boys who **perpetrated sexual AIPV (but not psychological)** were more accepting of sexual dating violence, affiliated with peers who they viewed as sexually abusive toward dating partners, and had not experienced psychological or physical AIPV in their dating relationships. 6. **Girls: More likely to report physical and psychological AIPV perpetration when**; they held attitudes that were more accepting of girls use of physical and psychological AIPV, perceived peers as being in abusive relationships and had been victims of psychological and physical abuse. **Were more likely to use psychological, and not physical AIPV if** they had experienced psychological AIPV and held less accepting attitudes towards physical AIPV.

Authors, (Year), Country	Type of Study	Study Aims	Participants	Exposure Assessments Used	Outcome Measures	Findings
Simon, Miller, Gorman-Smith, Orpinas and Sullivan (2010) North America	Cohort	To examine prevalence of **physical AIPV norms and prevalence of perpetration and victimisation of physical dating violence among early adolescents**	*N*=5404 6th Grade students selected from 37 schools in four U.S states. Male: 49% Female: 51% Age: 11-12 yrs. old. Ethnicity: 48% African American; 21% Latino; 18% Caucasian; 13% other or multiracial. 59.9% of males and 45.2% females in the sample reported having a boy or girlfriend in last 3 months (52% of overall sample).	Dating Violence Norms assessment: Adapted from Foshee et al. (1996). Measures beliefs about **girl's physical aggression towards boyfriends and then vice versa** in four parallel scenarios. Participants state whether they strongly agree or disagree (4-point scale).	Dating violence perpetration and victimisation: Adapted Foshee et al.'s (1996) existing **measure to assess physical violence perpetration and victimisation** Time scale- in the past 3 months.	1. 28.6% report **perpetrating physical AIPV** and 42.1% report **victimisation of physical AIPV** in past 3 months. 2**. Physical AIPV perpetration** was more common among females than males (31.4% vs. 26.4%) although those who perpetrated at the highest frequency (10 times or more) did not vary by sex. Male victimisation was higher (53.7% vs. 27.4%). 3. More accepted a girl hitting her boyfriend (52.9%) than a boy hitting his girlfriend (27.5%). **Attitudes supporting a girl-to-boy physical AIPV** was more prevalent among those who reported having a partner (59.8%) than boy-to-girl physical AIPV (32.3%) in the same sub-sample. 4. Those who held **attitudes supporting boys hitting girls** were significantly more likely to report physical AIPV perpetration (Adj. OR = 1.83; 95% CI 1.54-2.22), physical AIPV victimisation (Adj. OR = 1.70; 95% CI 1.42-2.04) or both (Adj. OR = 1.87; 95% CI 1.54-2.28). 5. Those who held attitudes supporting girls hitting boys were significantly more likely to report physical AIPV perpetration (Adj. OR = 1.83; 95% CI 1.50-2.22), physical AIPV victimisation (Adj. OR = 1.69; 95% CI 1.42-2.02) or as experiencing both (Adj. OR = 1.83; 95% CI 1.47-2.27).

Table 2.2

A summary of the quality assessment of the included studies

Author/ Year	Inclusion Bias	Selection Bias	Measurement/ Detection Bias	Attrition Bias	Outcome/ Result Bias	Statistics Used	Quality Assessment Score (%) *No. Unclear*
Ali, Swahn, & Hamburger (2011)	Parental consent obtained. Part of a larger Survey (Original N=4131). Sample selected had to have dated in past year (n=2888).	81% response rate in total survey. Variety of ages included. Similar ratio of male to female participants. Majority aged between 16-18 years. Difficulty in generalising to younger high school-aged students.	Self-report measures Attitudes assessed using a adapted tool. Alpha's for four constructs measured: Prescribed norms: .69 Perceived prevalence: .76 Gender stereotyping: .67 Perceived negative sanctions: .55. Scores dichotomised due to skewness problems. Excludes grey areas.	Those who had disability, required an interpreter or excluded or suspended (n = 353) ineligible.	Only **assessed physical violence perpetration and victimisation.** Adapted measure and included a number of items from the CTS. Reduced validity due to modification. Participants instructed not to include incidents they considered as self-defence. Good standard of internal consistency; Physical perpetration moderate, .92; Severe; .89 (identified from Foshee et al., 1996). A dichotomous outcome measure was created for each of the perpetration and victimization scales due to high skewness.	Descriptive Statistics Chi Square Analyses. Multivariate logistic regression analyses (Adj OR; 95% CI). Demographic characteristics and potential confounders were included in the analyses (i.e., personal competencies, problem behaviours, peer environment, family environment, and demographic characteristics. Also controlled for; grade level, sex, and race/ethnicity	44/58 (72%) *0 unclear*

Author/ Year	Inclusion Bias	Selection Bias	Measurement/ Detection Bias	Attrition Bias	Outcome/ Result Bias	Statistics Used	Quality Assessment Score (%) *No. Unclear*
Chase, Treboux, O'Leary and Strassberg (1998)	All participants had current/previous intimate partner and at least one same-sex peer whom they considered a friend. Parental Consent not clear.	Convenience Sampling from programme intake. Over-representation of males. Derived from high risk population. Cannot generalise to wider population.	Questionnaire developed for the study. Good internal consistency (Cronbach's alpha .74 to .89 for males; .80 to .96 for females). No additional validation studies weakening generalizability. Administered in small groups. Researcher ensured understanding by reading some of the vignettes out. Square-root transformation for skewness problems. Scores dichotomised.	No attrition rate reported.	CTS completed as part of larger battery of assessments. Unclear what these were. Standardized but reduced validity as **only used the physical aggression items** and original CTS was validated only for adult populations at time of study. Self-report States it explores the context of dating violence.	Descriptive Statistics Cochran Q tests Phi Correlations Point-biserial correlations Two-step Hierarchical regression	37/58 (64%) *2 unclear*
Feiring, Beblinger, Hoch-Espada and Haworth (2002)	22% of parents who were sent a consent form agreed child participation. Had to have	Under-representation of boys and younger high school students. Does include	Self-Report measures. Strong internal consistencies on measure; Dysfunctional Sexual Attitudes (α = .83); Tolerance for	68% did not respond to original parental consent form - lack of generalizability. 10% actively	**Measured perpetration and victimisation of physical violence (which included sexually coercive behaviours) and victimisation of emotional abuse.**	Descriptive statistics Canonical correlations MANOVA	40/58 *(70%)* *0 unclear*

Author/ Year	Inclusion Bias	Selection Bias	Measurement/ Detection Bias	Attrition Bias	Outcome/ Result Bias	Statistics Used	Quality Assessment Score (%) *No. Unclear*
	been in intimate relationship, currently or previously.	younger school age (11-12yrs). Cross-Sectional design does not allow for separating age from cohort effects or causal conclusions.	aggression (α = .86) and healthy relationship attitudes (α = .80). Validated though is still under development. Attachment questionnaire used was validated with good internal consistency. Internal consistency for modified version TOSCA-A was adequate.	refused to allow their child to participate. 3 students with parental consent refused to participate.	No measure of perpetration of emotional abuse due to time constraints in assessment. Participants asked for experience of dating violence in past year. Physical measure included sexual coercion item: "Kissed against will". Did not specifically measure sexual abusive behaviours which reduced validity.		
O'Keefe (1997)	Parents and students signed consent forms. All participants had to have been involved in a dating relationship.	Convenience Sample. Cannot generalise to non-school population. Good sample size Over representation of females.	Internal consistencies: Girl-to-boy violence items, α = .81 and .87 for boy-to-girl violence items. No further validation studies on adapted measure. No operational definition of justification attitudes.	2/8 high schools refused-belief that matter may be too sensitive. Low number declined participation or did not return consent form (20-25% in each class). 40 questionnaires	**Modified measure (CTS) used for both perpetration and victimisation of physical AIPV.** Cronbach's alpha for sample was .78. No time period stated so included 'have you ever inflicted…?' **Physical and forced sexual activity assessed**	Hierarchical Regression Analysis Pearson Correlation Descriptive Statistics	43/58 (74%) *1 unclear*

Author/ Year	Inclusion Bias	Selection Bias	Measurement/ Detection Bias	Attrition Bias	Outcome/ Result Bias	Statistics Used	Quality Assessment Score (%) *No. Unclear*
		Ethnically representative. 84% sample were aged 16-18 yrs.	No measure to explore other pro-violent attitudes. Self-report	eliminated as result of incorrect completion. 33 Questionnaires eliminated because they has not started dating.	**which may confound results. Psychological abuse not included.** Based on self-report. Standardized measure. Reduction in validity due to modification.		
Reeves and Orpinas (2012)	Consent obtained though unclear if this was from parents or participants or both. Part of a larger longitudinal study that followed a cohort from Grades 6-12. Current study was cross-sectional conducted when sample was in Grade 9.	Majority were randomly selected from larger study participants. 8% (n=52) were selected by teachers through being assessed as having difficulties with aggressive behaviour. Authors stated no differences found between these and those who were randomly sampled however potential for bias remains.	Good internal consistency: Four items measuring support for girls: Perpetrating physical violence against boyfriends (α = .85) and four items that measure support for boys aggressing girlfriends (α = .88). Confirmatory factor analysis used to test adapted scale structure.	Unclear whether all those who were originally selected completed the online surveys.	**Physical AIPV perpetration and victimisation**. Measure modified through reducing number of items and changing the time frame from 12 months to 3 months. Participants were instructed to exclude behaviours that they or their partner had committed in self-defence. Good internal consistency: .92 for perpetration and .91 for victimization measured by Cronbach's alpha.	Descriptive Statistics Multivariate logistic regression analyses; Fixed and Random Effects (Adj OR; 95% CI). Correlation (Pearson's *r*)	40/58 (69%) *2 unclear*

Author/ Year	Inclusion Bias	Selection Bias	Measurement/ Detection Bias	Attrition Bias	Outcome/ Result Bias	Statistics Used	Quality Assessment Score (%) *No. Unclear*
Sears, Byers & Price (2007)	Parental Consent Required. Had to have begun dating.	Sampling process not stated specifically. Students who expressed interest were given consent letter to take home. Possible selection bias. Reflected rural and urban areas. Lack of ability to generalise further to populations with differing characteristics. Demographic information not clearly reported. Younger age than other research (12 -13 yrs. lowest).	Self-Report. AWSA had good internal consistencies; α = .78 for boys and .75 for girls. Scale is validated. The six 'attitudes' scales reported good internal consistency and were validated tools. These were administered gender-specifically i.e. males only comments on attitudes to male-to-female violence and vice versa. No general pro-violent attitude scale. Administration of measures randomised which reduces practice effect.	190 excluded (71 girls, 118 boys, 1 did not state gender) due to having not started dating (*n*=148), exceeding the age range (*n*=14) or missing data (*n*=28) Unclear how many did not return consent forms or refused to participate as it states "very few".	**Self-Report of psychological, physical and sexual AIPV. Perpetration and victimisation.** CTS modified by reducing number of items. No reliability information provided. Participants presented with gender-specific versions. CTS Responses scored dichotomously. SES-R was modified to be gender inclusive and the original dichotomous score expanded to 3-point scale.	Descriptive Statistics Canonical correlation analyses	43/58 (74%) *1 unclear*

Author/ Year	Inclusion Bias	Selection Bias	Measurement/ Detection Bias	Attrition Bias	Outcome/ Result Bias	Statistics Used	Quality Assessment Score (%) *No. Unclear*
Simon, Miller, Gorman-Smith, Orpinas and Sullivan (2010)	Parental consent and student assent were obtained. All students, regardless of whether they were in a dating relationship commented on dating violence norms and a subgroup of those who dated were analysed further. Selected districts had higher poverty and youth crime rates than national USA average.	Random sampling from rosters from 37 schools. Two consecutive cohorts in sixth grade (aged 11-12yrs).	Adapted measure: good internal consistency for boys and girls yielding a Cronbach's alpha of .73 and .76 respectively. Eight items only (four items per gender). Scores dichotomised – excludes grey areas.	Consent rates ranged from 68%-84% across two cohorts and four sites. 9 student's data omitted due to patterned responses. 112 were deleted who were retained in sixth grade and randomly selected for inclusion in second cohort.	**Only assessed physical violence perpetration and victimisation.** Measure was adapted from an existing validated tool. Good internal consistency: $\alpha = .91$ perpetration and .89 for victimisation. Participants asked to exclude behaviours used in self-defence (reduced reliability of reports as no context/ motivation in incidents are known). Applying the 3 month time scale can limit responses/ exclude incidents. Scores dichotomised and categorised into experiencing both forms (perp/victim) or just one form.	Descriptive Statistics Multivariate logistic regression (adjusted Odds Ratio-OR at 95% confidence interval).	45/58 (78%) *0 unclear*

Results

The studies included within this review were those that examined the associations between violence-supportive attitudes and the different forms of AIPV perpetration and/or victimisation as a sole focus or as part of a multi-factorial analysis. Each of the study's results will be synthesised descriptively as opposed to statistically combined. This is due to the heterogeneity of the chosen samples, methodology, and findings.

Aims of included studies

Whilst all participants within each study were exposed to at least one assessment of attitudes relating to AIPV and compared on an outcome measure of self-reported AIPV perpetration, victimisation, or both; they differ in terms of the specific aims. Ali et al. (2011) and Reeves and Orpinas (2012) stated similar aims of exploring associations between attitudes supporting physical AIPV and experiences of physical AIPV perpetration and victimization although the latter study examined adolescents aged between 14 and 15 years old. Simon et al. (2010) examined this with early-aged adolescents (11-12 years old). Chase et al. (1998) explored the role of violence-justification attitudes in the specificity of age-mate relationships which included peer relationships, current dating partners and ex-partners. The remaining studies examined attitudes supporting AIPV as part of a multifactorial analysis to gain understanding of predictors associated with the infliction and victimisation of physical AIPV (O'Keefe, 1997) and in relation to specific types of AIPV tactics including physical, psychological and sexual AIPV (Sears, et al., 2007). The latter study's findings were presented specifically in relation to gender differences. Finally, Feiring et al. (2002) examined the relationship between dating violence attitudes (including acceptance of aggression, healthy relationship attitudes and sexual interaction attitudes) , physical AIPV perpetration and victimisation, psychological AIPV victimisation and the role of emotional and attachment styles. Despite

these differences in approach, the exposure and outcome measure were considered to reflect this review's aims.

Perpetration and Victimisation of AIPV

A number of the studies employed a modified version of the widely used Conflict Tactic Scale (CTS; Straus, 1979) as an outcome measure (Chase et al., 1998; O'Keefe, 1997; Sears et al., 2007). A further three studies used an adapted measure based on Foshee et al. (1996) (Ali et al., 2011; Reeves & Orpinas, 2012; Simon et al., 2010) which is reported to include a number of items from the CTS (Ali et al., 2011). Simon et al. (2010) also used the CTS to examine the physical aggression items relating to same-sex liked and disliked peers in addition to current and past dating partners. One study used an additional measure to identify motives of physical AIPV perpetration (O'Keefe, 1997). Sears et al. (2007) included psychological AIPV items within the CTS and used an additional measure examining sexually abusive behaviours towards dating partners.

All seven studies examined the prevalence of AIPV perpetration in some form. All but one study explored experiences of victimisation (Chase et al., 1998). One study (Ali et al., 2012) did not explicitly report of prevalence as this was examined in a previous study that used the same sample as part of a wider youth survey (Swahn et al., 2008). The majority of studies focussed on the physical perpetration of AIPV with psychological AIPV largely under-represented despite its most common prevalence (Leen et al., 2013). The finding that females reported inflicting more physical AIPV than males was consistent across all studies.

O'Keefe (1997) found 39% high school-aged males and 43% of high school-aged females self-reported perpetrating physical AIPV on at least one occasion. Whilst percentages indicating prevalence of victimisation were not presented, receiving physical AIPV was

found to be a significant (p = <.05) predictor in a hierarchal regression model of AIPV perpetration for both males and females. This shows support for bi-directionality (Haynie et al., 2013; Leen et al., 2013). Further support for this was found in Feiring et al. (2002). Whilst the occurrence of physically aggressive behaviours within dating relationships were not common (female perpetration: 29%; male perpetration: 15%; female victimisation: 23%; male victimisation: 15%), over half of the sample (N = 254) reported victimisation of emotional abuse. The prevalence increased among students aged 16-18 years old. There were significant gender differences, with more females perpetrating physical AIPV than males (χ^2 (1, N = 254) = 6.1, $p \leq .01$). Whilst females self-reported a higher level of physical victimisation, this did not reach significance.

Upon examining students aged 14-15 years, Reeves and Orpinas (2011) found that one-fourth of their sample who dated reported perpetrating at least one act of physical AIPV and one-third reported victimisation (n =404 males and females in total). More females reported perpetration (33%) and more males reported victimisation (37%). There were strong correlations between perpetration and victimisation for both males and females, .78 and .75 respectively (p =.01 for both) highlighting reciprocal engagement in AIPV. Simon et al. (2010) found a high prevalence of physical AIPV perpetration amongst younger adolescents in their study of 5,404 students in sixth grade (aged 11-12 years). Out of 2,806 students who reported having a girl/boyfriend, almost one-third of females and one-fourth of males (31.5% versus 26.4%) reported being physically aggressive towards them. These results demonstrated a significant gender difference. However, this study did not dichotomise physical AIPV perpetration as others did and found that there were no significant sex differences between males and females who reported the highest frequency (5.6% for males and females). The results further demonstrates bi-directionality with 77% of perpetrators reporting victimisation and 52% of victims reported AIPV perpetration.

In a sample of 'high-risk' adolescents aged between 14 and 19 years old, Chase et al. (1998) found high levels of physical aggression across all relationships. The self-reported prevalence of physical AIPV perpetration against current partners was 33% of males and 68% of females. With regards to specificity, males' use of aggression occurred as part of a more generalised pattern of violence within dating relationships (towards current partners and ex-partners) which did not expand to same-sex peers. Whereas females use of physical AIPV was partner specific and unrelated to aggressive behaviour in other relationships. Ali et al.'s (2012) study also utilised a sample of 'high-risk' adolescents examining physical AIPV however prevalence of AIPV was not commented separately from its interactions with attitudes supporting AIPV. This information is available in Swahn et al. (2008).

Sears et al. (2007) was the only study to investigate the prevalence of different types of AIPV perpetration and victimisation; physical, psychological (emotional) and sexually abusive behaviours. In a sample of 633 high-school students, they found that 43% of males and 51% of females had reported engaging in at least one form of AIPV. Results revealed 19% of males and 26% of females engaged in more than one form of AIPV. Sears et al. (2007) commented that a number of studies had excluded the use of psychological AIPV which may lead to underestimation of prevalence. Within this study, males' use of psychologically abusive behaviour was the most common (35%) with physically abusive and sexually abusive behaviour being used by 15% and 5% of the male sample respectively.

Attitudes condoning the use of AIPV

The included studies varied with regards to the self-report measures utilised to explore attitudes supporting AIPV. Three studies (Ali et al., 2011; Reeves and Orpinas, 2012; Simon et al., 2010) applied modified AIPV attitude measures used by Foshee and colleagues in previous studies (1996, 2001, 2005). One study created an attitudinal measure for the purpose

of the study (Chase et al., 1998) and O'Keefe (1997) used a modified 'justification of violence attitudes' scale used previously by Margolin and Foo (1992). Feiring et al. (2002) utilised a previously developed measure which included items examining dysfunctional sexual attitudes and healthy relationship attitudes in addition to items exploring tolerance of aggression. Sears et al. (2007) applied a series of measures that had been developed by Price et al. (1999) which measured justification of AIPV within the specific abusive behaviour domains, i.e. a measure for each form of AIPV. Further, Sears and colleagues (2007) administered a measure examining beliefs towards women.

Generally, studies found low levels of supporting attitudes for each form of AIPV (Feiring et al., 2002; Reeves and Orpinas, 2012; Sears et al., 2007) and liberal attitudes towards gender roles (Sears et al., 2007). Interestingly, the younger aged students held high levels of attitudes supportive of physical AIPV (Feiring et al., 2002; Simon et al., 2010) and lower healthy relationship attitudes compared to those in their late-teens (Feiring et al., 2002). Those who utilised 'high-risk' samples found high levels of attitudes justifying physical AIPV (Ali et al., 2011; Chase et al., 1998).

In the majority of the cases, there was a gendered direction in attitudes held whereby males held higher levels of justification or acceptance for male-to-female directed physical AIPV and females held such beliefs for female-to-male directed physical AIPV (Reeves and Orpinas, 2012; Simon et al., 2010). It was evident that males held a more accepting view of sexually abusive behaviour within relationships (Sears et al., 2007). The authors suggested that this may be part of a gender role script and may view such behaviours as normative as opposed to abusive, however the methodology did not allow for this to be analysed. Notably, though perhaps unsurprisingly, female-to-male aggression was more accepted amongst participants highlighting a 'double standard' (O'Keefe, 1997; Reeves and Orpinas, 2012;

Simon et al., 2010) although Reeves and Orpinas (2012) stated that physical dating violence was not largely accepted by participants.

Associations between attitudes condoning AIPV and AIPV perpetration and victimisation

Using a regression model, O'Keefe (1997) found that justification of physical AIPV in different situations was a significant predictor of physical dating violence infliction whilst highlighting gender difference. Males were more likely to inflict violence against a dating partner when they believed male-to-female violence was justifiable ($r = .42$, $p < .05$). For females, the infliction of dating violence was linked to the belief that female-to-male violence was acceptable ($r = .22$, $p < .001$) but when male-to-female violence was not. This study also identified a number of other factors related to the infliction of dating violence including substance misuse, history of aggression, relationship seriousness, school violence, parent-child violence and witnessing inter-parental violence. The sample used here has been described as 'high-risk' (Ali et al., 2011) although this was not specifically stated by O'Keefe. The socioeconomic status of the overall sample was highlighted as below the national average which has been identified as a risk factor associated with AIPV perpetration and victimisation (Ali et al., 2011).

Within their high-risk sample Ali et al. (2011) also found a significant association between physical AIPV perpetration and victimisation and attitudes supporting physical AIPV using a cross-sectional logistic regression model. The results remained significant after controlling for psychosocial factors, high-risk behaviours, demographics and peer and family influence. Gender-directed associations were identified whereby attitudes condoning male-to-female violence were linked with males' physical AIPV perpetration and victimisation and

the tolerance of female perpetration was statistically associated with females engaging in physical AIPV. Further, this attitudinal acceptance of female-to-male violence was associated with physical AIPV victimisation for both boys and girls. Similarly, Reeves and Orpinas (2012) identified a strong association between acceptance of male-to-female violence and males' physical AIPV perpetration and victimisation. The role of attitudes amongst females and their experience of physical AIPV were either weakly associated or not significantly associated.

Consistent with the above findings, Simon et al. (2010) found strong associations between involvement in physical AIPV (perpetration or victimisation) and acceptance of physical AIPV attitudes among early adolescents. Those who held attitudes supporting male-to-female hitting were significantly more likely to report AIPV perpetration (Adj. OR = 1.83; 95% CI 1.54-2.22), victimisation (Adj. OR = 1.70; 95% CI 1.42-2.04) or as experiencing both (Adj. OR = 1.87; 95% CI 1.54-2.28).

Sears et al.'s (2007) results found that males' and females' use of multiple forms of AIPV were predicted by their attitudes towards and experiences with violence. Although attitudes condoning different forms of AIPV were not common amongst the sample, they found that males who held more accepting attitudes towards the use of each form of dating violence and more gender traditional roles of women were more likely to report having used psychologically, sexually and physically abusive behaviour in their dating relationships. However, holding traditional views was not common amongst participants. More specifically, boys who were sexually abusive but not psychologically abusive were more accepting of sexual AIPV and perceived that peers were engaging in sexually abusive behaviour towards their dating partners.

In exploring the specificity of AIPV, Chase et al.'s (1998) results provided modest support. They found that justification for aggression within relationships was predictive of violence within specific partnerships. For males, justification of physical aggression towards a previous dating partner significantly predicted dating aggression against a current partner, $F(1, 59) = 7.32$, $p < .01$, $R = .33$. For females, there was a moderate correlation between their justification and use of aggression towards a current partner, but not for violence towards a previous partner. In summary, dating aggression was found to be relationship-specific for males and partner-specific for females.

In contrast to the above studies, Feiring et al. (2002) found no significant relationship between tolerance of aggression and physical AIPV perpetration or victimisation nor psychological AIPV victimisation. However, results indicate that physical AIPV perpetration was positively related to dysfunctional sexual attitudes and negatively to healthy relationship attitudes. Victimisation was not associated with the attitudes measured within the study. Stronger relationships were found between emotional styles such as externalising responsibility for the harm caused to others. The interpretation of this result may arguably be linked to beliefs relating to perceived consequences of dating violence. The authors state that the lack of significant findings was contrary to expectations.

Within the studies that controlled for potential confounding variables, several were significantly associated with perpetration of mainly physical AIPV including same-sex peer violence perpetration and victimisation (Ali et al., 2011); low self-efficacy (Ali et al., 2011); AIPV among peers (Ali et al., 2011; Sears et al., 2007); witnessing inter-parental violence (O'Keefe, 1997) or fearing family violence (Sears et al., 2007).

Discussion

Key Findings

This review examined seven studies in relation to the association between attitudes among adolescents towards at least one form of AIPV and the infliction or victimisation of at least one form of AIPV. The study samples examined within this review varied between 95 and 5,404 with age ranging from 11-20 years old. The aim of this review was to examine:

Hypothesis 1

1. The extent of the association of pro-violent, acceptance and/or justification attitudes with perpetration and/or victimisation of each type of AIPV; physical, sexual and psychological/emotional.

Through examining the results of each study, this review largely demonstrated that there are modest to strong concurrent associations between attitudes that justify or support the use of the different forms of AIPV perpetration and victimisation which is consistent with the findings of Leen et al. (2013), however such attitudes were not overly common amongst community samples (Sears et al., 2007). This relationship appears more strongly amongst high-risk samples (Ali et al., 2011; O'Keefe, 1997) and younger-aged adolescents (Simon et al., 2010). Furthermore, this association was found in one study to be relationship-type specific for males and partner-specific for females (Chase et al., (1998). On the contrary, Feiring et al. (2002) did not find any association between tolerance of aggression and AIPV. The implications of these findings are discussed in the section below.

Table 2.3 outlines the studies that supported/ partially supported the presence of these associations. It is evident throughout the current review that the included studies were overwhelmingly focussed upon physical perpetration and victimisation. Just one study (Sears et al. 2007) examined all forms of AIPV within their sample thus rendering it impossible to

make generalisable conclusions regarding the extent of association between attitudes and psychological and sexual AIPV. Given the lack of specificity within these studies, no conclusion could be made concerning whether attitudes are more strongly associated with one form of AIPV compared to another. These findings must be considered within the context of differing methodology and focus. The limitations and implications of this will be discussed below.

Table 2.3

The extent to which studies found association between attitudes and AIPV perpetration and/or victimisation

Association found	Partial association found
O'Keefe (1997)	Chase et al. (1998)
Sears et al. (2007)	Feiring et al. (2002)
Simon et al. (2010)	
Ali et al. (2011)	
Reeves and Orpinas (2012)	

Hypothesis 2

> 2. Whether attitudes vary depending on the specific categories of AIPV perpetration and/or victimisation, including physical, sexual and psychological/emotional abuse.

Similar to above, the second research question in this review requires further investigation before firm conclusions can be made regarding psychological and sexual AIPV. This is due to the majority of studies utilising a measure that examined attitudes that support or justify physical AIPV only, i.e. 'is violence ok if…' (Ali et al., 2011; Chase et al., 1998; O'Keefe et

al., 1997; Reeves & Orpinas, 2012; Simon et al., 2007). The exception was the study of Feiring et al. (2002) who examined various attitude types including tolerance of aggression, dysfunctional sexual attitudes and healthy relationship attitudes within the context of physical AIPV perpetration and victimisation, and emotional AIPV victimisation. Their findings highlight the importance of considering a broader array of attitudes within studies of AIPV, and indeed adult IPV, particularly because of the implications that physical AIPV may not be associated to an individual's 'tolerance' or 'acceptance' of aggression per se, but more related to dysfunctional attitudes and misunderstandings of what constitutes a healthy relationship (Callahan et al., 2003).

Sears et al. (2007) was the sole study in this review which utilised a measure examining attitudes of acceptability relating to males use and females use of each form of AIPV; physical, sexual and psychological perpetration and/or victimisation. They found that acceptance attitudes of all types of AIPV behaviour predicted all types of dating violence in boys which does not provide support for different attitudes distinguishing between different types of AIPV. However, Sears' et al.'s (2007) canonical correlation analysis identified distinguishing features associated with the use of sexually abusive behaviour and the absence of psychologically abusive behaviour. This included acceptance of sexually abusive behaviour by males towards females and the perception that peers were engaging in sexually abusive behaviours. This may tentatively indicate that specific attitudes may encourage specific behaviours within APV and requires further investigation.

Of those studies that did explore a wider array of attitudes (towards acceptability, gender roles, dysfunctional sexual attitudes and healthy relationships) there were some differences found in the strength of association with AIPV. However the results provided limited insight due to the lack of other studies available for comparison that have explicitly examined attitudinal specificity and its relationship to the various forms of AIPV. Findings will be

discussed in light of the aforementioned limitations so as to not overstate the findings and enable interpretation along with recommendations for future directions in research.

Interpretation of the findings

Perpetration and Victimisation of AIPV

All but one study examined the prevalence of both perpetration and victimisation of AIPV which could strengthen conclusions among a community sample of adolescents. The prevalence of total AIPV perpetration was found to be between 15% - 39% for males and 26% - 68% for females with higher levels found amongst high risk samples. On the whole, similar rates of around one third of the samples were identified. The results concur with Leen et al. (2013) with regards to more females reporting the perpetration of physical AIPV. Understanding the prevalence of AIPV, and adult IPV, is highly important in informing the level of provision required for primary, secondary and tertiary interventions. This indicates the need for prevention programmes to continue developing using the findings from methodologically robust studies.

There is a lack of a standardised operational definition of AIPV used in this area of research and this is apparent among the studies within this review which impacts the ability to draw firm conclusions. There were several variations of the included acts that constitute dating violence which has implications for the accuracy of reported perpetration and/or victimisation and therefore the extent of the association with attitudes condoning such acts. As highlighted previously, the majority of studies examined physical AIPV only (Ali et al., 2011; Chase et al., 1998; Reeves & Orpinas, 2012; Simon et al., 2010). O'Keefe (1997) and Feiring et al. (2002) examined physical violence but the items within the outcome measure included sexual coercion items which impacts on the validity of the results. Further, Feiring

et al. (2002) did include psychological abusive behaviour but within the context of victimisation only. Sears et al. (2007) explored different types of dating violence separately which enhances the measurement validity of the results. The exclusion/inclusion of particular types of dating violence in this current review is inevitably impacted by the continued application of different working definitions (Barter et al., 2009). Although some argue for the exclusion of sexual AIPV (Glass et al., 2003), the CDC (2008) explicitly state physical, sexual and emotional abuse constitutes AIPV which enables the adoption of consistent methodology within research studies. Inconsistency and exclusion of psychological and sexual AIPV has implications for the clinical significance of the findings, particularly in the ability to generalise to other forms of dating violence outside of the identified behaviours within each study. The notable absence of the studies examining psychological AIPV, which has been identified as the most common form of AIPV (Leen et al., 2013) highlights further need for studies to include psychological, and sexual, AIPV specifically as a means of a more robust understanding of the AIPV paradigm. It is suggested that future studies should provide a rationale for the inclusion/exclusion of specific types of AIPV.

Given that the current research is focussed predominantly upon physical violence, so too are AIPV interventions (Safe Dates Program, Foshee et al., 1996). Whilst attitudes towards violence are addressed in these programmes, this may often be within the context of the acceptability and justification of hitting or pushing a partner. The predominant use of psychological AIPV found within research should be utilised in practice whereby participants are not just educated as to the differing types of AIPV but also encouraged to consider, question and address the attitudes and behaviours in the context of psychological AIPV. This also applies to sexual AIPV perpetration and victimisation where addressing attitudes relating to expectations of sexual intimacy within a relationship and healthy relationship attitudes should be included in both prevention and intervention.

The outcome tool to measure AIPV behaviours also has implications when interpreting the results. Whilst internal consistency was reported as satisfactory in the majority of studies, this information was unavailable in others (i.e. Sears et al., 2007) despite adjustments being made to the measures. The modifications within the measures limit the ability to make direct comparisons to other studies measuring the construct of AIPV. During the time of the earlier studies (O'Keefe, 1997; Chase et al., 1998), the CTS had been validated for adult samples and not for adolescents. Later studies employed measures that had undergone validation research (Foshee et al., 1996; Price et al., 1999).

Given the self-reported nature of the measures, individuals may have been reluctant to report the true extent of AIPV relating to the social connotations attached. The issue of under-reporting of abusive behaviour within intimate relationships is widely referenced (see Chapters one and three. Some studies asked participants to restrict their self-reporting of behaviours to those which had occurred within the last three months (Reeves & Orpinas, 2012; Simon et al., 2010), others identified 12 months (Ali et al., 2011; Chase et al., 1998; Feiring et al., 2002) and two studies had no specified timeframe. Restricting the time-frame could contribute to under-representation of reported AIPV perpetration and victimisation.

A finding of AIPV bi-directionality was observed within the studies along with some gender differences. In line with the typology of adult IPV perpetrators (Holtzworth-Munroe, 2000; Johnson, 2006), identifying the characteristics of individuals who are 'perpetrator only', 'victim only' or 'both' can provide clinically meaningful information when considering criminogenic need and in the development of appropriate intervention. Some studies requested exclusion of behaviours that they or their partner used in self-defence (e.g. Simon et al., 2010) which has an impact upon the clinical significance of the findings regarding further insight into the contextual factors associated with AIPV. Furthermore, three studies dichotomised their outcome variable results (i.e. perpetration: yes/no; victimisation;

yes/no). All authors reported that this was due to skewedness of findings however it must be noted that dichotomising results may reduce the ability to distinguish between the rarer occurrences of severe and frequent experiences of AIPV amongst a sample. The majority of typology research has thus far been concerned with adult IPV perpetrators (Johnson, 1995; 2006) and lack of empirical evidence to date may provide reasons as to why such methodology is not used in AIPV studies. However, it has recently been found that Johnson's (1995; 2006) typology provides a useful framework in understanding AIPV (Zweig, Yahner, Dank & Lachman, 2014). Adopting this approach may provide an important insight into risk and criminogenic need among adolescent-aged perpetrators of dating violence. In the academic field, it will enable more robust methodologies (Foshee, Bauman, Linder, Rice & Wilcher, 2007). In terms of clinical practice and intervention, it is suggested that programme facilitators consider the different typologies when delivering AIPV programmes to ensure a more responsive approach to each participant.

Attitudes and AIPV

A variety of different assessments for measuring attitudes towards dating violence were utilised across the studies, with no two studies using the same measure as described in the Results section above. The majority of the tools demonstrated good internal consistency although the measure utilised in Ali et al. (2011) ranged from questionable to good (.59-.76). Administration of more than one measure may be viewed as a more in-depth assessment of attitudes.

Overall, this review did find consistent associations between acceptance of AIPV attitudes and experience of AIPV (predominantly physical) although one was in relation to physical AIPV and dysfunctional sexual attitudes and healthy relationship attitudes (Feiring

et al., 2002). What was evident, particularly amongst younger samples, was the seeming trend in which female-to-male aggression was more accepted amongst participants highlighting a 'double standard' (O'Keefe, 1997; Reeves and Orpinas, 2012; Simon et al., 2010). This finding within the current review highlights the need for a clear move away from the more traditional approaches of targeting patriarchal attitudes of violence towards women within AIPV intervention and the importance of examining and changing more general violent attitudes.

Studies did not consistently examine the role of such attitudes for each type of AIPV, creating difficulties for conclusions in relation to the specific aims of this review. This association must not be overstated as the studies also outlined a range of other factors associated with AIPV perpetration and/or victimisation. For example, O'Keefe (1997) highlighted a number of predictors associated with AIPV perpetration and victimisation which included witnessing inter-parental violence. This finding places recognition of the impact that adult IPV may have upon children who witness it. Research relating to gender differences within the intergenerational cycle of IPV may provide further understanding of the findings that females use of AIPV tended to be partner-specific as opposed to them engaging in AIPV across all intimate partnerships (Stith et al., 2000).

Each study must be considered within the context of measurement difficulties. Measuring attitudes among populations is historically problematic given its subjective nature. This leaves construct validity open to question particularly where assessments have not been standardised or validated. Consistency was evident in that all studies employed a measurement of attitudes relating to the acceptance, condoning or justification of dating violence however the format varied. Ali et al. (2011), Feiring et al. (2002), Reeves and Orpinas (2012) and Sears et al. (2007) used measures that included general statements about whether AIPV was deemed acceptable within dating relationships. The interchangeability of

the terms ‘acceptance’ and ‘justifications’ was apparent within the studies and it could be argued that these constructs are fundamentally different when attempting to gain detailed understanding of the AIPV paradigm. This makes comparisons difficult. For example, a ‘justification’ of AIPV may refer to whether an act of AIPV is perceived as warranted within a specific contextual situation whereas an ‘acceptance’ or tolerance of AIPV could be viewed as an individual holding more general AIPV-supportive beliefs. Chase et al.’s (1998) study is supportive of the latter statement although it would be beneficial for further studies to examine this in order to demonstrate consistency of findings amongst various adolescent samples.

Since the relationship between attitudes and dating violence is not always clear (Leen et al., 2013; See chapters three and four), more sophisticated questions must be asked such as the “why?” and the “how?” when studying this link (Schumacher & Smith Slep, 2004). Measures should aim to examine social norm attitudes in more depth as opposed to simplifying the construct to “acceptable versus not-acceptable”. In terms of implications for intervention, effort to unravel the context and triggers to ‘acceptance’ or ‘justification’ attitudes is required by facilitators to help participants understand these processes. A fuller understanding for both participant and facilitator can serve to increase the effectiveness of subsequent work to address such attitudes.

O’Keefe (1997) approached their study from a social learning theory perspective to explore the context of AIPV incidents which is encouraging in terms of research applying theory-practice links to further enhance understanding. Only three of the seven studies (Chase et al., 1998; O’Keefe, 1997; Simon et al., 2010) administered measures that examined the contextual justification of AIPV using vignette examples of behaviours. These included whether a person justified AIPV in situations involving jealousy, betrayal, embarrassment or when feelings a lack of support from their partner. Whilst a vignette may not reflect

responses based upon what an individual would do themselves in reality, it does provide a context in which participants reported that the use of AIPV was justified. Unfortunately, results from these measures were presented as an overall score as opposed vignette specific results. Such information is crucial for understanding the specific interactions between attitudes and behaviour. Further, O'Keefe's (1997) research was the only study that explored motives of AIPV however did not explicitly examine or comment upon the possible relationship with attitudes towards AIPV. Should such links be explored within this study, and others, the findings could have been further examined utilising cognitive dissonance theory (Festinger, 1957) and hostile attribution bias (Dodge et al., 1990).

Other methodological considerations

The design of all selected studies was non-experimental meaning that no comparison between dating violence couples and non-dating violence couples could be made. As such, it is not possible to generalise the findings to all adolescents within dating relationships or ascertain whether pro-violent attitudes are higher among those who have experiences of dating violence in comparison to a population who have not.

As the studies are cross-sectional in nature, no cause-effect inferences can be made, i.e. the direction of the relationship. Whilst this review's aim was to examine to concurrent relationship between attitudes and specific types of AIPV, this design leaves the results open to differing interpretations that may be bias to a researcher's preferred theoretical approach. It could be possible that an individual could employ an attitude that justifies their use of aggression after an incident of dating violence has occurred as a form of rationalisation and reducing cognitive dissonance (Festinger, 1957).

An increasing number of longitudinal studies have been carried out within this field of research (Leen et al., 2013; see Vagi et al., 2013) and have highlighted some inconsistent results pertaining to the stability of the predictive strength of attitudes upon AIPV perpetration and victimisation. Given the increased validity of experimental study designs, this area of research would benefit from further studies that adopt more rigorous experimental approaches and standardised operational definitions of AIPV.

With regards to age, five out of the seven studies captured several age groups and two studies examined a specific age group (Reeves & Orpinas, 2012; Simon et al., 2010). This has implications for the interpretation and generalisability of results pertaining to participants' engagement, interest and experiences of dating relationships. Adolescence is a period of significant development with significant dating relationships being considered to occur between the ages of 15 and 16 years old on average (Bethke & DeJoy, 1993). However, nearly half of Simon et al.'s (2010) sample reported having a girlfriend or boyfriend in the past three months despite the younger age range of that stated above. The implications are such that a 12 year-olds experience may differ significantly in relation to an 18 year-olds experience with regards to the seriousness and emotional significance they place upon that relationship. The finding that adolescents aged 11 to 12 years hold a high rate of attitudes that are accepting of violence highlight the need for more research to examine differences in attitudes in early-aged adolescence.

The ages of those included in the population was raised to 21 years old to include individuals whom had been held back in high school. This impacted upon one study, and the specific number of this age in the sample was unobtainable. Despite this, the mean age throughout all studies was no more than 17 years. As 21 years of age exceed the age of "adolescence", this may affect the applicability of the findings to the population of interest.

Notably, the issue of sexual orientation within each of the studies was rarely commented upon. Only one study (Reeves & Orpinas, 2012) stated that the majority of their sample consisted of heterosexual participants although six had reported dating same-sex partners or both sexes. Due to the limited number, these individuals were included amongst the rest of the sample however any implications of differing sexual orientation were absent from discussion.

The majority of the study samples were recruited by convenience sampling which may enhance the possibility of bias although six out of the seven studies had what would be considered large sample sizes (Field, 2009). Linked with this is the attrition bias for those who did not obtain parental consent and could not take part in the studies. Non-completing participants or those that refused to engage may have particular characteristics that identify them as a higher risk for being perpetrators or victims of domestic violence, such as having lower socioeconomic status, experiencing parental violence or having lower school attainment (Vagi et al., 2013).

All included studies explicitly state their limitations and identify the need for more rigorous and creative longitudinal designs.

Strengths and weaknesses of the current review

There are a number of limitations that need to be considered in relation to the method utilised within this systematic review. This has implications for the clarity of the findings. Whilst a pre-defined inclusion/ exclusion protocol was adopted in an attempt to reduce bias, it does not eliminate bias completely. Origins of bias in this review include publication bias, as time constraints restricted the search strategy to only three electronic databases. Ideally, several other electronic databases would have been searched and those that were not available would

have been obtained as a hard copy. Further publication bias is evident as unpublished journals were not included.

The pre-defined inclusion/exclusion criteria were relatively strict. Articles that were not in English were excluded due to a lack of resources for translation, introducing language bias and may have implications for the potential mediating effect of cultural influences on dating violence attitudes. The lack of consistency between the exposure and outcome measures employed limited the ability of this review to make confident conclusions. All seven studies relied wholly on self-report in both the exposure and the outcome measures. Given the sensitive nature of the topic, it is likely that the results were affected by a level of social desirability in responding and an under-estimation of reporting of perpetration of domestic violence. This concept is also relevant when examining individuals' reports of victimisation. Furthermore, the studies did not include couples' reports despite the increasing evidence for mutual AIPV (Vagi et al., 2013). This may have increased the cross-validity of self-reporting and thus the quality of the included studies.

In terms of sampling bias, no UK studies were identified as suitable for the current review which reduces the generalisability of the findings to a UK sample. Studies that researched samples of adolescents in custody or psychiatric hospitals were excluded as the current review was aimed to reflect those within the community. This has implications for identifying those who may be at further risk of developing pro-violent attitudes or inflicting (or being a victim of) dating violence. A review examining this specific population would be beneficial in identifying population differences.

Participants in five out of the seven studies had to have been in a previous or current dating relationship to be included. Whilst individuals who have not been involved in a dating relationship cannot report incidences of abusive behaviours, the inclusion bias may impact

the results in terms of having a higher base line of individuals who hold accepting views towards violence. This is supported by research linking the experience of violence, vicarious learning and attitudes (Bandura, 1986).

Given the non-experimental design of the studies, the quality scores were relatively low. When analysing the quality of the statistical tests used, possible bias may arise given the lack of access to the original data. The author was the only assessor in applying the quality assessment procedure, making it impossible to assess inter-rater reliability. Such methodology would enhance the validity of the findings. The findings of the studies should be considered within the limitations of generalisability and not to inform trends outside of the samples investigated or amongst a wider adolescent population.

Conclusions and Implications

In terms of the specific aims of this review, it has not been possible to draw firm conclusions. Whilst it provides evidence that attitudes condoning AIPV is modestly-strongly associated with physical AIPV perpetration and victimisation, it is not possible to draw clear conclusions about the specific role attitudes play in psychological or sexual perpetration and/or victimisation of dating violence or the whether specific attitudes are more commonly associated with different forms of AIPV. Gender differences were not examined in detail due to the focus of the current thesis upon male perpetrators of IPV and AIPV however the consistent differences highlight a continued need for further research.

It is apparent that multiple layers within the construct of attitudes are associated with the experience of dating violence including; differing attitudes such as justification, acceptance of general violence, perception of gender roles, healthy relationship attitudes and affiliating with peers who hold similar views. O'Keefe (1997) and Sears et al. (2007)

identified a number of other significant predictors, including demographic, contextual and situational factors. This reinforces the complex nature of AIPV and in terms of intervention; a consideration must be given to this in both development and delivery.

It is also evident that more research needs to be conducted in the areas of psychological AIPV, sexual AIPV and the role of specific attitudes in relation to this. Further examination of the role of attitudes in relation to the following areas: abusive behaviour typologies (including among one-sided and mutually violent couples); the context within which AIPV occurs (in relation to jealousy, desire for control, impulsive anger); dyadic variables and attitudes examining 'victim', 'perpetrator' or' both' categories (Stith et al., 2012); victim/ perpetrator specific attitudes, and further comparison studies using control groups.

The complexity of the role of attitudes in AIPV has implications for AIPV intervention development and delivery. Fellmeth et al.'s (2013) review of programme effectiveness demonstrated that changes to a participant's attitudes that condone violence were marginal, similar to findings in adult intervention research (Babcock et al., 2004; see Chapter three). This may be reflective of low base line levels of attitudes condoning violence amongst participants (Feiring et al., 2002; Reeves & Orpinas, 2011) or indeed suggest that interventions are not currently targeting such attitudes effectively. Continued research into 'why' such outcomes occur is required. The consideration of the suggestions made above relating to interventions above may assist programme developers and practitioners in asking further questions relating to the function of attitudes in AIPV. The benefits of this will be that participant and practitioner will gain a deeper understanding of the context of behaviours and place the role of social cognition amongst a range of other risk factors associated with AIPV. More research of an experimental and longitudinal design is required to gather a deeper understanding and to generate inferences pertaining to causality. Furthermore, it will serve to

enhance the generation of theories that are evidence-based. This will also have significant implications of the development and implementation of effective prevention and intervention programs to address dating violence.

Chapter Three

Empirical Research Study: An exploratory study into the effectiveness of the Community Domestic Violence Programme (CDVP) within a Probation sample of adult male perpetrators.

Chapter Three Rationale

As discussed in Chapters one and two, the prevalence of IPV perpetration at adolescent- and adult-age remains a concern within society and effective tertiary interventions are required as part of addressing this social problem. There is a view from some scholars that interventions aimed at IPV offenders appear to have been less influenced by 'what works' literature in their development in comparison to other offender populations such as sexual offenders (Hanson et al., 2009) and are considered to take an ideological approach to IPV (Bowen, 2011; Dixon et al., 2012). There is a need to continue to evaluate IPV intervention programmes empirically using non-biased theoretical approaches to progress forward within the field.

Abstract

This study examines the effectiveness of IPV treatment utilising pre-, post- and six-month post-treatment psychometric scores and recidivism data. Furthermore, it aimed to compare treatment completers, non-completers and recidivism data amongst those who completed the programme. Data was collected using a sample of 259 males; 216 treatment completers and 43 non-completers who had been court ordered to attend the Community Domestic Violence Programme (CDVP) in the community. Data were obtained from Thames Valley Probation Trust and statistically analysed. Results indicated that completion of CDVP had a positive impact upon participants' self-reported psychometric scores, yielding moderate to large effect sizes. Notably, treatment completers self-reported an increased ability to manage their

assessed anger, a reduced level of experiencing anger routinely (Trait anger), a reduction in feeling like expressing anger verbally or physically and a reduction in reported levels of jealousy across a number of the domains, excluding sexual jealousy.

Low levels of IPV-related recidivism (reconviction and alleged IPV re-offending) were observed for both treatment completers and non-completers. Non-completers were significantly more likely to be convicted of an IPV offence however this difference disappeared when IPV reconviction and IPV alleged re-offending were aggregated to expand the construct of recidivism. Treatment completer recidivists were more likely to be younger and surprisingly reported slightly lower levels of avoidant attachment at the post-treatment stage. Non-completers were categorised as a higher level of risk, were significantly more likely to breach their order and reported higher levels of jealousy (threat to exclusivity) and a higher avoidant attachment style.

As this thesis aims to examine factors relating IPV across the lifespan, the discussion will include literature linking adult and adolescent interventions and consider practice issues. The findings are also discussed within the context of the study's methodological limitations.

Introduction

The Risk Need Responsivity (RNR) model developed by Andrews, Bonta and Hogue (1990) posits that offender treatment programmes are more effective when they consider; the level of programme intensity to an offender's risk level (risk principle); target the relevant criminogenic needs of an offender (need principle); and when the modality of the intervention reflects an individual's learning style and abilities (responsivity principle; Andrews, Bonta & Wormith, 2011). The RNR model provides principles that underpin the development of such programmes, allowing for the interpretation of the offender assessment and treatment literature and is viewed as a term which is now synonymous with the effective

risk management of offenders (Ward, Rose & Willis, 2011). This has provided an evidence base that the rehabilitation of offenders does 'work' when treatment approaches adhere to the RNR principles (Andrews & Bonta, 2010a; 2010b; Andrews et al., 2011; McGuire, 2013; Ward, Melser & Yates, 2007).

Walker, Bowen and Brown (2013) comment that the current framework of IPV treatment programmes has not progressed in the same manner of other offence-specific interventions such as sex offender treatment (Hanson, Bourgon, Helmus & Hodgson, 2009). Further, it has been argued that interventions aimed at IPV offenders appear to have been less influenced by the RNR principles and the 'what works' literature in their development and take an ideological approach to IPV (Bowen, 2011; Farmer & Callan, 2012). Despite this, feminist scholars continue to debate the benefits and innovation of the Duluth model. Indeed the vast majority of intervention programmes do utilise a multi-modal approach employing cognitive behavioural techniques alongside the psycho-educational element (Babcock et al., 2004; Hamilton et al., 2013).

There has been an increase in evaluating IPV treatment programmes however in recent years, its effectiveness in reducing re-offending remains inconclusive (Davis & Taylor, 1999; Eckhardt, et al., 2013; Gordon & Moriarty, 2003; Ministry of Justice, 2010; Rosenfeld, 1992). Duluth-based interventions that focus on the role of patriarchy as a central criminogenic need continued to be used despite theoretical criticism (See chapter one; Dixon & Graham-Kevan, 2011; Dutton & Corvo, 2006; Farmer & Callan, 2012). However there is evidence that no treatment modality is better than another within IPV interventions, i.e. cognitive-behavioural versus pro-feminist (Akoensi, Koehler, Lösel & Humphreys, 2013; Babcock et al., 2004; Hamilton et al., 2013). It is suggested that these theoretical debates relating to the incompatible nature of the feminist versus the gender-inclusive viewpoint may provide more conflict in theory than in practice due to combined-approach programmes

making distinction more difficult (Hamilton et al., 2013; Stover, Meadows & Kaufman, 2009). What is clear is that these debates have made exploring the effectiveness of IPV perpetrator treatment programmes continually difficult with many studies and reviews presenting different results (Feder et al., 2011). It is likely that these differences have been impacted by a lack of consensus relating to the purposes and aims of IPV interventions and treatment modality (Hamilton et al., 2013). Indeed, it is explicit that all IPV interventions ultimately aim to reduce IPV perpetration. However, the use of various indicators of 'treatment effectiveness' can lead to a heterogeneous outcome evaluation (Hamilton et al., 2013) thus impairing the ability to make consistent conclusions and comparisons (Bowen, 2011).

Treatment evaluation reviews among heterosexual male IPV perpetrators have predominantly been conducted in North America with a review of Europe and the UK treatment appearing more recently (Akoensi et al., 2013). Discussion relating to female perpetrators and those who perpetrate IPV within a same-sex relationship is beyond the scope of this paper given its focus on heterosexual males (see Baker et al., 2013; Carney & Dutton, 2007; Thornton, Graham-Kevan & Archer, 2012).

When utilising recidivism at a treatment outcome, Rosenfeld (1992) found that an average of 27% treatment completers reported recidivism on at least one occasion (weighted by sample size in 25 studies). This statistic reduced to between 5 and 13% when police data was used and increased to 36% when spouse self-report was used highlighting the issues with under-reporting of 'official figures' (Tjarden & Thoennes, 2000) with victim/partner reports being regarded as a more 'accurate' outcome measure (Feder & Wilson, 2005). There was little information regarding drop-outs rates within the studies that examined court mandated intervention programmes. Rosenfeld (1992) concluded that completion of IPV treatment had

no significant effects upon recidivism indicating other explanatory factors are involved in cessation of offending.

On the contrary, Davis and Taylor (1999) found that the mean effect size of quasi-experimental studies (treated participants versus untreated participants) was .416 and .412 for the experimental studies, showing 'fairly consistent evidence that treatment works' (Davis & Taylor, 1999, p. 69). Babcock et al.'s meta-analysis (2004) concluded that IPV interventions had only a small yet significant effect upon recidivism. They argued that the 5% difference in chance of non-violence (40% non-violence for treatment completers compared to 35% non-violence for untreated offenders) could equate to 42,000 less women being victimised in the United States. Akoensi et al. (2013) highlighted that the effects were reduced when Babcock et al. applied more methodologically stringent criteria to the included studies.

Feder and Wilson's (2005) meta-analysis, based on ten North Amercian studies, initially found a reduction from 20% to 13% recidivism following treatment based on official reports. However, this finding disappeared when partner reports were used as an outcome measure. When examining studies that employed quasi-experimental designs with treatment drop-outs as the comparison group, they found that the random effects mean effect size was $d = .97$, $p \leq .05$. When fixed effects were used, this reduced to $d = .49$, [95% CI of 0.27-0.71], but remained significant. The significance of the random effects analysis could indicate that the findings were generalizable beyond these quasi-experimental studies used in this meta-analysis. However, the researchers suggest that such findings cannot be attributed to the programme due to possible confounding variables associated with lack of programme attendance and group differences. It is evident that more experimental studies yield smaller treatment effects highlighting the need for further rigorous studies.

Stover et al. (2009) conducted a meta-analysis on the available IPV treatment studies in North America that used randomized controls with a sample of 20 or more in each group.

They reported an overall recidivism rate of 30% within six months of those who had completed treatment regardless of the treatment modality (Duluth versus CBT-based). The authors raised concerns relating to high levels of drop-outs within the programmes (up to 30%) and overall study attrition rates of victim reports in the follow-up periods (15-89%). This led to caution when making any solid conclusions about the effectiveness of perpetrator programmes. The lowest recidivism rate (18%) was found when perpetrators engaged in couples-treatment that simultaneously addressed substance misuse and issues with aggression. This demonstrated preliminary support for couples-treatment and such findings promote the need to take a wider approach to IPV treatment given the co-morbidity of issues associated with IPV perpetration (Gilchrist et al., 2003; O'Leary et al., 2007). This is supported by Eckhardt et al. (2013) who found equivocal results when it came to the effectiveness of traditional IPV intervention programmes, but promising results relating to more innovative treatments.

It is argued that the lack of meta-analyses within this area reflects the relatively few methodologically robust IPV outcome studies (see Bowen, 2011). Methodological difficulties have been widely documented as a reason for the lack of consistent results across IPV treatment effectiveness domain as a whole (Feder et al., 2011; Gondolf, 2011). More specifically, Saunders (2008) highlighted that most studies in the IPV literature are focused on recidivism rates and generally, these studies are plagued with high attrition rates, poor controls, lack of structured programming, differing definitions of recidivism, small sample sizes, and inconsistent or limited follow-up periods. Indeed, evaluating treatment effectiveness is wider than solely considering recidivism (Laws & Ward, 2011; Walker et al., 2013) and despite its challenges; continued research is required to enhance knowledge about effective practices within IPV treatment.

More recently, a two part review of IPV treatment practices in Europe was published (Akoensi et al., 2013; Hamilton et al., 2013) aiming to address the surprising dearth in the literature relating to UK studies. Akoensi et al. (2013) provided a systematic literature review (Part 2) on 12 studies that utilised both attitudinal and behavioural outcomes. The intervention programmes used different theoretical approaches, similar to that observed in North America. Some positive outcomes were found within the studies relating to self-reported psychological change however these were again critiqued owing to limited methodological rigour which led to inconclusive results with no estimated effect size.

Shorey and colleagues highlighted the lack of success to date in efforts to reduce dating aggression amongst adolescents (2012). Group-based intervention programmes for adolescents are currently being evaluated in North America (Foshee et al., 2005). Whilst there are cautiously positive outcomes these largely reflect the modest results associated with adult IPV interventions. However, Foshee and colleagues (2005) have found that positive behavioural changes have been observed in areas such as possessing weapons and less-peer violence although there remains a dominant focus upon attitudinal change in such evaluations (Leen et al., 2013). There remains a dearth of these studies within the UK and despite forward movements occurring within this area, robust longitudinal evaluative research is required to inform treatment efficacy.

The above highlights the challenges associated with IPV interventions however it is important to acknowledge the increasing work that is being conducted to enhance the understanding of IPV treatment effectiveness. Gondolf (2011) suggested that conclusions from IPV treatment evaluations could be unduly pessimistic due to the studies not always examining the full impact of treatment. For example, there has been a progression in the conducting and publishing of treatment effectiveness studies in the UK (Bowen, 2010; Bowen, 2011; Bowen & Gilchrist, 2006; Bowen, Gilchrist & Beech, 2005; 2008).

A holistic approach to evaluating treatment effectiveness is imperative to enhance understanding (Bowen & Gilchrist, 2004). This includes examining both attitudinal and behavioural variables. Bowen et al. (2008) examined individual change within a community based pro-feminist psycho-educational IPV intervention programme. Psychological characteristics (such as pro-violent attitudes, locus of control, interpersonal dependency and anger) were investigated in addition to re-offending data obtained from police records. The sample consisted of 52 offenders and 32 non-offenders as a comparison group. The authors concluded that whilst half of the offenders achieved significant psychological change, this was considered as limited and did not relate to re-offending. This demonstrates individual differences and the need to further examine *why* one modality of treatment may benefit one person and not another. Further, this study highlighted the complex relationship between psychological changes versus behavioural change. Bowen et al. (2005) found that higher levels of interpersonal dependency and more involvement with the police prior to starting the programme predicted post-treatment recidivism. This raised the need to further explore the possibility of sub-types within treatment effectiveness studies (Straus, 2008; Walker et al., 2013).

The above studies were based upon a feminist model prior to the implementation of accredited treatment programmes within the UK. As far as the author is aware, there has yet to be a published evaluation paper relating to the Community Domestic Violence Programme. This programme utilises a predominantly cognitive-behavioural approach and is accredited within UK's Probation Trusts and HM Prison Service.

Current IPV Interventions within the UK

There are two main accredited intervention programmes used within the UK: the Integrated Domestic Abuse Programme (IDAP) and the Community Domestic Violence Programme

(CDVP). The former is based mainly upon the Duluth Model (Pence & Paymar, 1993) although utilises Dutton's nested ecological model (1995; 2006) to a certain extent. The latter is based upon the Family Violence Perpetration Programmes from the Correctional Services in Canada (CSC) which places more focus upon the multi-faceted explanations of IPV and has Dutton's nested ecological model at its core.

IDAP is delivered in the majority of Probation Trusts within the England and Wales, with CDVP being delivered in fewer areas (33 areas versus nine; Dixon et al., 2012). However, CDVP is also delivered within HM Prison Service where it is known as the Healthy Relationships Programme (HRP). There are two levels of intensity of the HRP, medium and high. The medium intensity programme is identical to CDVP whereas the high intensity programme targets higher risk offenders reflecting the 'risk' principle of increased treatment provision and intensity (Andrews et al., 1990). Both of these programmes are specifically targeted at adult male perpetrators and include work with known victims and inter-agency risk management.

The CDVP aims to decrease; distorted beliefs that negatively influence the appraisal of a trigger, emotional mismanagement (in relation to anger, jealousy, fear and dependency), problems in self-regulation and coping and social skills deficits in order to increase an offender's ability to forge healthier relationships (CDVP Manual, 2008). Participants are encouraged to critically analyse and change their thinking patterns with the aim of reducing abusive and violent behaviour within their intimate relationships and families. Further, the programme aims to increase the offender's ability to respond non-abusively, to empathise with victim(s) and give offenders a greater sense of personal responsibility for their violence. This is achieved using a range of psychological theories and techniques in addition to the use of cognitive behavioural strategies including; rational emotive behaviour therapy, solution-focused therapy, cognitive therapy, motivational interviewing, psycho-educational

approaches and a relapse prevention framework to help increase the offender's ability to identify and manage high-risk situations. It may be argued that the CDVP's focus upon various criminogenic needs is more reflective of the current 'what works' literature relating to risk factors linked to IPV perpetration (see Chapter one) in comparison to sole feminist-based approaches.

The CDVP is based upon the Medium Intensity Family Violence Programme (MI-FVPP) and the High Intensity Family Violence Programme (HI-FVPP; that which HRP is based upon). These programmes were evaluated in 2005 (Stewart, Gabora, Kropp & Lee, 2005). Participants had completed the programme whilst in custody and follow-up took place within the community. Stewart et al. (2005) found significant self-reported improvement in all attitudinal measures, including decreased jealousy and abusive attitudes, increased responsibility and empathy and a reduction in reported IPV for programme completers versus non-completers. In the high-intensity group, significant differences were found in recidivism rates for treatment completers in comparison to untreated offenders, 4% versus 14% respectively (significant at .05). These differences were not statistically significant in the moderate-intensity group, but using odds ratio the findings suggested that untreated offenders were 3.25 times more likely than treated offenders to commit spousal violence, and 1.57 times more likely to commit any violence. It may be difficult to generalise the findings of this study to community completers given that incarcerated offenders may experience less situational triggers upon which they can realistically be tested in terms of their attitudinal and behavioural change. More recently, Connors, Mills and Gray (2012) evaluated the MI-FVPP on incarcerated offenders and found significant improvement across all aspects of the psychometric battery ($p < .001$) demonstrating positive associations between treatment completion and self-reported healthier functioning. The researchers, who used a within-between design, found that higher motivation as perceived by the group facilitators was

associated with higher reported psychological improvement. The attrition rate was 15.8% of the total programme and they highlighted the importance of motivation amongst drop-outs and as a possible means to enhance the positive effects of IPV treatment. Limitations of this study included the lack of a randomised control group and reliance of self-report, although this was attempted to be countered by incorporating facilitator ratings of motivation levels.

The above findings related to motivation have been considered to be integral in the effectiveness of treatment in other literature (Ward & Stewart, 2003) with RNR scholars recognising the need for a "renewed focus on offender motivation" (Andrews et al., 2011, p. 750). This is imperative given IPV drop-out rates can be 40-50% among the IPV perpetrator population (Bowen & Gilchrist, 2006). McMurran and Theodosi (2007) found that profiles of non-completers show that they are more high risk than completers and are more likely to re-offend, as did Olver, Stockdale and Wormith (2011).

Understanding of IPV is evolving and broadening (Stith et al., 2012). The studies by Stewart et al. (2005) and Connors et al. (2012) provided inspiration for the current study. The author aimed to evaluate the transferred MI-FVPP programme among a UK setting in the form of CDVP to further inform understanding of the effectiveness of IPV interventions within the context of community offenders.

Aims and Hypotheses

The aim of the present study is to examine the effectiveness of IPV intervention among adult male offenders who have completed CDVP comparing psychometric results and recidivism. This study compares the treatment effectiveness of completers and non-completers of the CDVP programme. The alternative hypotheses for this study are:

H1: There will be a significant difference in pre-treatment scores for treatment completers and non-completers.

H2: There will be a significant difference in attitude changes among treatment completers (pre, post and six-month post-treatment).

H3: There will be significant differences between treatment completers and non-completers in relation to re-offending rates.

H4: There will be a significant difference among treatment completer recidivists and non-recidivists in relation to their demographics and post-psychometric scores.

Method

Sample

A 'post hoc' quasi-experimental cross-sectional design was used within this study with the analysis of data occurring at one point in time. The sample was selected retrospectively from a database used within the Thames Valley Probation Trust identifying people who had completed, or commenced but not completed, CDVP between April 2011 and November 2012. This was to enable the collection of psychometric data at the six-month post-treatment stage. Participants were selected after September 2009 due to a change in the standard battery of administered psychometrics. The original sample consisted of a total of 307 offenders. From this number, 48 were excluded from the study due to lack of access to information, leaving a total sample of 259. The mean age for the whole sample was 35.51 years (*SD* = 10.67 years; range = 20-72 years). The majority of the sample (80.3%) was described as White whilst 6.9% were described as Black, 4.2% as Mixed, 3.1% as Asian and 0.8% described as Other which included those of Chinese ethnicity (4.6% did not disclose). 52.9% were employed at the time of referral to the programme.

The majority of the sample were assessed as medium risk of harm (91.8%) and 8.1% considered high risk of harm to a known adult using the Offender Assessment System (OASys, Howard, 2006; See Materials section for further details). Whilst 36.7% had a

previous violent conviction, 89.2% had a history of IPV offending (convicted or not convicted). Of those with a history of IPV (n = 230), 30.4% had previous conviction(s) for an IPV-related offence, 23.5% had an IPV- related caution(s), 7.4% received both a conviction(s) and caution(s) and 34.4% of the sample self-reported varying forms of IPV perpetration prior to the commencement of CDVP within their current or previous intimate relationships. Of the sample 68% were subject to a Community Order, 29% to a Suspended Sentence Order and 3.1% were required to complete the programme as part of their Licence conditions. All participants were mandated to complete CDVP as part of their Order. At the time of the offences, 25.5% were in a co-habiting relationship and 18.5% were married. 29.7% had separated from their victim prior to the offence. Out of 195 individuals who had children or step-children (75.3% of sample), at least one child was present during the offence in 68.7% of cases (i.e. 134 offences). Further, 72.2% of the sample were under the influence of substances at the time of the offence (alcohol = 58.3%; drugs = 5.8%; both = 8.1%) with 47.9% having mental health problems either at the time of the offence or in their history. This included evidence of referral to a GP, receiving medication or counselling, or self-reporting difficulties relating to depression, anxiety or suicidal ideation. Levels of stress were also considered to represent mental health difficulties if an individual had suffered a nervous breakdown or had contact with professional services relating to this issue.

The sample was split into two groups who had been convicted of an offence against an intimate partner; Treatment completers (n =216, M = 32.74 years, SD = 10.11, Range = 20 – 55 years) and non-completers (n = 43, M = 36.06 years, SD = 10.71, Range = 20 – 72 years). The age of both treatment completers, $D(216) = 0.098$, $p < .001$, and non-completer's, $D(43) = 0.148$, $p < .05$ were statistically non-normal. As such, a Mann-Whitney U was conducted which confirmed no significant difference between age (U = 3774, z = -1.94, p =

.052, r = -.12). The non-completers group make up 16.6% of the total sample. See 'Description of Sample' for further detail of group differences.

Ethical Considerations

Prior to conducting the research, ethical approval was gained from the University's Science, Technology, Engineering and Mathematics Ethical Review Committee and the National Offender Management Service (NOMS) which included the Head of Public Protection, Thames Valley Probation Trust and the Head of Thames Valley Probation Trust's CDVP Programmes Team.

Written consent was obtained for all participants prior to conducting the study. Upon enrolment of the CDVP, individuals sign a consent form which states that they consent to the requirements of the programme and the questionnaire data to be used to enable the evaluation of individual changes and the programme's effectiveness (see Appendix six). This form is fully explained to group members' prior to signing.

The sample included some offenders who were under supervision by the Probation Trust at the time of data collection and also offenders whose supervision orders have ceased. No participants were directly contacted during the study due to its retrospective nature and no data was obtained about participants once their order had ceased, unless they had re-offended and sentenced. This study analysed archived data from April 2011 to November 2012. It was therefore not appropriate to contact offenders to debrief the study as they may no longer be under statutory supervision.

The extracted data and all reported findings were anonymised with participants assigned a number to enable the tracking of treatment completion, psychometric completion and recidivism outcomes.

Variables

Participant information varied marginally depending upon whether an individual was a treatment completer or a non-completer (see Appendix seven). Information examining participants' recidivism rates following completion or non-completion of CDVP was also obtained (see Appendix eight). To ensure consistent data extraction over time, a pro-forma was utilised (see Appendix nine). Ideally two researchers would have extracted data to ensure inter-rater reliability however this was not possible and is recognised as a limitation of this study.

The Community Domestic Violence Programme (CDVP)

CDVP is an IPV treatment programme delivered in Thames Valley Probation Trust. It is designed for heterosexual adult males (aged 18 years and over) who have received a conviction of perpetrating violence against a current or ex-partner. It also includes offenders convicted of offences that stop short of physical violence but have the same intention (e.g. harassment, criminal damage, threats or malicious calls). Those suitable must be identified as medium or high risk of harm as identified from OASys or the Spousal Abuse Risk Assessment (SARA, Kropp et al., 2005). The programme consists of 26-28 sequential core group sessions comprising six modules of differing session lengths; Module One: Motivation and Enhancement; Module Two: Awareness and Education; Module Three: Managing Thoughts and Emotions related to Abuse; Module Four: Social Skills; Module Five; Relapse Management and Module Six: Healthy Relationships. Psychometric questionnaires are completed in a pre-group and a post-group session. Each session lasts up to two hours 15 minutes. It should be noted that details of the programme delivery here is specific to Thames Valley Probation Trust and may vary between Trusts. The Offender Manager (OM; probation

officer) is an integral part and are involved in pre-group sessions, meetings during the programme and after completion to enhance motivation and consolidation of learning.

Programme Integrity

A review of the implementation of accredited domestic violence programmes (both IDAP and CDVP) was completed by the Ministry of Justice (Bullock, Sarre, Tarling & Wilkinson, 2010). This focussed largely upon the logistics of programme implementation, effectiveness of multi-agency working, programme integrity and group members' experiences of the programmes. Bullock et al. (2010) found that the material for each programme was delivered within the framework with deviation occurring only in response to group dynamics or issues (responsivity principle). Within the Thames Valley Programmes Team, every session was recorded and video monitoring of sessions was conducted a minimum of three times by a trained treatment manager. Regular supervision and video monitoring feedback was scheduled into each programme.

Measures

The psychometric measures used within CDVP are nationally distributed to all offenders enrolling in the programme. The aim of each psychometric assessment is to measure change in areas that have been empirically identified as relevant in the perpetration of IPV. Four questionnaires are administered at the pre-treatment, post-treatment and six-month post-treatment stages.

The State-Trait Anger Expression Inventory Second Edition (STAXI-2; Spielberger, 1999)

The STAXI-2 is widely used across a range of populations and cultures. It contains a 57-item inventory comprising of three scales: State Anger (S-Ang: to measure the intensity of anger as an emotional state in a specific situation); Trait Anger (T-Ang: to measure an individual's disposition to experience angry feelings routinely), and Anger Expression and Control (AX/AC: to measure anger expression and control). For accurate interpretation of the STAXI-2, percentiles and/or T scores are required (Spielberger, 1999). For the current study sample, corresponding scale and subscale percentile ranks for normal adults was used to indicate how an individual compares to with others' of a similar age/gender. For full details and a critique of the measure, see Chapter four in this thesis.

Interpersonal Relationship Scale (IRS: Hupka & Rusch, 2001)

The IRS is a 27-item likert scale that assesses six aspects of romantic jealousy: Threat to exclusive relationships, Envy/self-deprecation, Dependency, Sexual possessiveness, Competition and vindictiveness, and Distrust (Hupka & Rusch, 1989; 1992; 2001). Responses are made on a six-point scale (ranging from 1 = strongly agree to 6 = strongly disagree) with lower scores indicating higher levels of jealousy. Subscale scores can be obtained by summing up the respective item scores. A total score can be obtained by adding all item responses although this was unable to be obtained in the data set. There are no norms available for the IRS in general (Personal Communication, R Hupka, 28th May 2013) however the NOMS Test Battery Guide (2012) provided norms using a UK sample of IPV offenders (See Appendix ten). Internal consistency (Cronbach's alpha) on each of the six subscales using the same sample of UK IPV offenders are reported as: Threat to exclusive companionship, $\alpha = .85$ ($N = 9417$); Self-deprecation/envy, $\alpha = .82$ ($N = 9353$); Dependency, $\alpha = .86$ ($N = 9577$); Sexual possessiveness, $\alpha = .52$ ($N = 9486$); Competition and vindictiveness, $\alpha = .47$ ($N = 9611$); Distrust, $\alpha = .65$ ($N = 9551$). A positive treatment effect

would represent a higher score when compared to the pre-treatment score (Stewart et al., 2005). IPV perpetrators have been found to hold higher levels of fear of abandonment and the presence of pathological jealousy, particularly sexual jealousy (Fernandez-Montalvo, Echeburua & Amor, 2005). It is notable that this scale was abandoned by the developer's (Personal Communication, R Hupka, 19th October 2013) however normed data for UK IPV offenders demonstrate its utility amongst them.

The Experience in Close Relationships – Revised Questionnaire (ECR-R, Fraley, Waller & Brennan, 2000)

The ECR-R is a 36 item tool designed to assess individual differences with respect to two underlying attachment-related patterns: Anxiety and Avoidance. The anxiety dimension refers to the extent to which people are insecure vs. secure about the responsiveness and availability of their partner. The avoidant dimension represents the extent to which people are uncomfortable being close to and depending on others. Respondent's rate the items on a seven-point Likert scale (1 = low anxiety/avoidance to 7 = high anxiety/avoidance). The results are plotted onto a two dimensional continuum to identify the interaction between attachment-related anxiety and avoidance. A low score on both anxiety and avoidance suggests a secure attachment.

The normed population for the ECR-R was a sample of 22,000 individuals made up of 78% women which should be taken into consideration when considering the results of the current study. The means of the normed males were Anxiety: M = 3.64 and Avoidance: M = 2.88. The internal consistency reliability tends to be .90 or higher for the two ECR-R scales (Sibley & Liu, 2004) although lower at the secure end of both dimensions (Fraley et al., 2000). The correlation between the avoidance and anxiety scales was .41. Test-retest reliability was .94 for Anxiety items and .95 for Avoidance items (Fraley et al., 2000). The

NOMS Test Battery Guide (2012) demonstrate the internal consistency of the ECR-R amongst UK IPV offenders as .91 (N = 3893) and .85 (N = 3863) for Anxiety and Avoidance, respectively. Avoidant attachment style is more positively correlated to IPV than anxious attachment in some studies (Doumas, Pearson, Elgin, & McKinley, 2008). Although, Follingstad et al., (2002) found that anxious attachment in early childhood resulted in the development of an angry temperament, which in turn led to a controlling style and physical aggression within an adolescent sample.

The Revised Attitudes Towards Offence Scale (RATOS)

The RATOS is a 37-item five-point likert scale self-report questionnaire with zero representing strongly disagree to four representing strongly agree. It has been developed by combining elements from existing tools and extended to include items relating to denial, minimisation, blaming and responsibility due to an absence of an existing validated measure of attitudes towards IPV offences specifically. This measure is under development (NOMS Test Battery Guide, 2012). Scores above 50 indicate the presence of denial, minimisation, victim-blaming and failure to take responsibility, at least to some extent. Scores of 75 or above indicate firmly held views. There is currently no evidence of validation against other measures and should therefore be used as a measure for assessed attitude change as opposed to a precise diagnostic tool. Despite the lack of validation, it has been used within this study as it is routinely used at the pre- and post-treatment stages of CDVP. Norms have been provided by NOMS (NOMS Test Battery Guide, 2012) based upon 704 IDAP and CDVP participants: Pre-programme stage is M = 41.22 (SD = 12.85) and M = 27.65 (SD = 13.11) at the post-programme stage. For men, minimization of conflict and partner blame were associated with self-reported perpetration of intimate partner aggression, even after

controlling for socially desirable responding and relationship satisfaction (Scott & Straus, 2007).

The psychometric information was obtained using IAPS. The data was available in raw score format and converted according to percentile rank for the STAXI-2 and an average score for the ECR-R. The raw scores were obtained and used for the IRS and RATOS.

The Offender Assessment System (OASys: Howard, 2006)

OASys is an IT based structured clinical judgement assessment tool used within correctional services in England and Wales. Professionals completing OASys are able to link an individual's offending-related needs to the risk of serious harm throughout the assessment and identify a level of risk: Low, Medium, High or Very High. The OASys manual identifies risk of serious harm as "…a risk which is life-threatening and/or traumatic, and from which recovery, whether physical or psychological, can be expected to be difficult or impossible" (cited in Howard, 2006, p.2).

Definition and Measuring Recidivism

Recidivism is generally defined as "a falling back or relapse into prior criminal habits, especially after punishment" (Blumstein & Larson, 1971). It is noted that studies examining recidivism use a wide range of definitions which impact the outcome and interpretation of results (Saunders, 2008) and the difficulties in measuring recidivism. Whilst it is important to recognise that recidivism is not ideally used as a sole outcome measure, such studies can inform policy and programme development by informing researchers of how many people recidivate, the characteristics of such individuals, and the frequency and nature of the behaviours (Maltz, 2001). Within this study, individuals were considered to be recidivists if they had been convicted of a non-IPV related offence or an IPV offence, had received a police call-out following a domestic dispute or a police caution or reprimand. Further

information was considered in relation to Order breaches following programme suspension or completion of CDVP. No time scale was available for this.

Offender self-reports was also utilised and obtained through the Integrated Case Management System (ICMS) database that is used within the Probation Service. All entries on ICMS were completed by the OM. Other methods included reports from checks with the Domestic Abuse Unit (DAU).

The follow up period was identified using the treatment completion or treatment suspension date, the order end date and date of first offence or incident for each participant. This difference was calculated into months as this is most common quantity used in recidivism studies.

Procedure

All information was obtained electronically using OASys, ICMS, the Interim Accredited Programme Software database (IAPS) and a Microsoft Excel spreadsheet database used by the Programmes Team which held demographic, enrolment, completion and suspension information. Participants were selected using an opportunity sampling method and split into two groups: 1) Treatment completers and 2) Non-completers. Non-completers comprise of individuals who began the programme but did not complete it at any time nor engaged in any type of subsequent IPV offending behaviour treatment. This enabled comparisons to be made within and between the groups to enhance the methodological robustness of the study.

The data collection for recidivism was completed in August 2013 allowing for a maximum follow up period of up to 28 months. The follow-up period in this study ranged from one month to 22 months. A set follow-up period was not obtained due to the method of data collection (i.e. information on participants was not available once they had completed their Order unless they had re-offended).

For each participant, information was gained on a number of different variables. The type of offences committed for the current IPV conviction, IPV reconviction/re-offending and general reconviction were categorised through offence type including offences against property, offensive behaviour, harassment, violent offences and drug offences. The coding differed slightly between general reconviction and IPV-related reconviction/re-offending due to the nature of the offences. Where an individual had more than one offence, the most serious offence was coded using guidance from Schedule 15 of the Criminal Justice Act 2003 (CJA, 2003, Sch 15). In most cases, this was violent offences against a person. For the current IPV conviction, the offences were also separated into further categories as many had more than one offence: contact, non-contact or both (contact and non-contact). Contact offences included physical violence ranging from battery to murder. Non-contact offences included harassment, threats to kill, possession of a weapon and criminal damage within the context of a relationship. Where a participant's current offence included a physical assault as well as non-contact act behaviour, they were classified in the 'both' category. Each category was distinct and did not overlap. A further category of 'other' was included for 'relationship to the victim'. This accounted for victims such as acquaintances or family members of the individual whom the perpetrator had an intimate relationship with.

Information was obtained on for the reason of suspension for the non-completer group to enable further insight into drop-out characteristics. The Breach variable identified individuals who had breached their order at any point from completing the programme (or suspension date) until their Order End Date. A further offence may result in a breach if an individual has been resentenced consecutively. In other cases an individual may not receive a breach in response to a new offence. For IPV offenders, it is standard procedure that the OM gains periodic updates from the DAU. In some cases, the OM's had not received DAU information or it was unavailable to be viewed by the researcher (the information was stored

in a paper file). DAU information was available for 43.6 % of the sample. Self-report to the OM was the main source of information as recorded on the ICMS database where this was unavailable. In five cases (two treatment completers; three non-completers), the individual self-reported an incident that the DAU had recorded to their OM. Where this occurred, it was recorded under DAU alone. This information was only obtained for the 'alleged IPV re-offence' category (where no conviction occurred) because conviction information was available through OASys.

In relation to recidivism data, each group (treatment completer and non-completer) was split into further categories based upon whether they had been re-convicted or re-offended: 'Non-IPV reconviction' (referring to a general offence conviction), 'IPV re-conviction' and 'alleged IPV re-offence'. The latter included self-reported IPV incidents, DAU call-outs or receiving police cautions. This was in order to widen the conceptualisation of 'recidivism' given the under-reporting of IPV (Hester, 2006). Where further incidents or offences were identified for a participant, details were gained by examining the participant's OASys.

The independent variables (that which represent the input or cause of an effect) within this study include CDVP at the pre, post and six-month post-treatment stage, demographic variables and the participant groups of 'treatment completers' and 'non-completers'. The dependent variables (which represent an effect or outcome) include the psychometric scales and subscales and recidivism rates.

Analysis

Data was analysed using SPSS Statistics Software V21. Initial tests were carried out to examine data assumptions. Kolmogrov-Smirnov (K-S) tests were conducted to ascertain the normality of data distribution. K-S tests were conducted on treatment completers and non-

completers and amongst recidivists and non-recidivists within the treatment completer group. This was carried out on each group separately as it is important to consider the distribution within each group as opposed to the distribution overall (Field, 2009, p. 147). Data is said to significantly differ from a normal distribution if the *D* statistic (K-S test) is < .05. The 'treatment completer' group consisted of 216 participants. Given that the K-S test is notoriously affected by large samples in which small deviations from the normality yield significant results, the significance level was increased to $p < .01$ as opposed to the standard $p < .05$ limits (Field, 2009, p. 139). In cases where a variable was not normally distributed in the majority of conditions but found to be normally distributed in one condition, it was viewed as not meeting the assumptions of normality and non-parametric test was utilised.

A series of Levene's tests (Levene, 1960) were carried out to identify homogeneity of variance (the variance of the outcome variable should remain consistent at all levels of the predictor variable) which compared the two groups for any significant differences in variance. This test is less important with large sample sizes due to Central Limit Theorem which assumes that in a large sample size the sampling distribution will be normal (Field, 2009). However, homogeneity of variance becomes more of an issue where there are unequal sample sizes which needed to be considered when conducting between-group analyses (Field, 2013). The K-S test identified that the majority of variables were not normally distributed (more often than not as a result of skewness). Therefore, the *F* statistic in the Levene's test was conducted using the median for variables that were not normally distributed as this is found to increase the robustness and power of the Levene's test against skewness (Brown-Forsythe Test; Brown & Forsythe, 1974).

Pearson's Chi-Square (χ^2; Pearson, 1900) analyses were conducted to explore any significant associations between the treatment completers and non-completers and recidivists and non-recidivists. The assumptions include that the data must be dependent (each value

must contribute to one cell only) and the expected frequencies should be greater than five as then χ^2 loses statistical power. It is reported as being acceptable within larger sample sizes to have up to 20% of expected frequencies below five (Field, 2009). The Fisher's exact test (FET) was used to explore association between recidivists and non-recidivists due to the smaller sample size and to adjust for expected frequencies lower than five (Fisher, 1922). Cramer's *V* Statistic provided further information as to the strength of any significant association: Under .10 = Negligible association, .10 and under .20 = Weak association, .20 and under .40 = moderate association, .40 and under .60 = relatively strong association, .60 and under .80 = strong association, .80 to 1.00 = very strong association (Rea & Parker, 1992, p. 203). Whilst data was obtained on both IPV and non-IPV offending, statistical analyses were focussed upon IPV recidivism data due to the nature of the study.

Mann-Whitney U (Mann & Whitney, 1947) and t-tests were conducted based upon whether the data met parametric assumptions. This was first conducted on pre-treatment scores for treatment completer's and non-completers, and recidivists and non-recidivists, to identify any differences (higher or lower) in attitudes or criminogenic need. A Bonferroni correction was applied to minimise the chance of Type 1 error given the numerous paired tests that were conducted.

Despite allowing at least six months following completion, the attrition rate of the six-month post-treatment questionnaires was very high (76% attrition). The six-month post-treatment dependent variable consisted of 50 people in comparison to original 216 representing 23% of the original treatment completer sample. This impacts on the power of the findings in relation to the respective hypothesis. Attrition was due to offenders completing their order prior to the psychometric due date, death of offender (in one case), failure to complete, being remanded or refusal to complete the psychometrics.

A series of Friedman ANOVA's were utilised to compare attitudinal change among treatment completers at the pre, post and six-month post stages where data were not normally distributed. A MANOVA cannot be conducted unless data is normally distributed and ANOVA is said to be robust to unequal sample size where all other assumptions are met (Field, 2009). Therefore, a one-way related-sample ANOVA was used in circumstances where the data met parametric assumptions. Post-hoc analyses were conducted where appropriate.

A Life Table Survival Analysis was used to examine rates of recidivist's survival over time. Survival analysis is a statistical method for the study of time to an event (Mahmood, 2013). This is a non-parametric statistical test as it makes no assumptions about the nature of the data and the survival function being estimated (Garson, 2012). A Life Table provides a description of the distribution of time-to-an-event variables and the probability from each of the intervals are estimated. In this case, the event was 'Overall IPV recidivism'. IPV reconviction was not examined separately using this procedure due to the small number of those reconvicted within the sample.

Ferguson (2009) stated that 'increasing emphasis has been placed on the use of effect size reporting in the analysis of social science data' (p. 532) therefore effect sizes are presented throughout this study. This provides clinical meaning to findings by estimating the level of change. In this study, appropriate effect sizes were reported based upon the statistical test employed. This included Pearson's correlation coefficient *r* (Pearson, 1900) which is considered an appropriate effect size for non-parametric tests such as the Mann-Whitney U test (Field, 2009). Pearson's *r* may be interpreted as .10 = small effect, .30 = medium effect and .05 = large effect. Cohen's *d* was used as it is viewed as a parametric effect size (Cohen, 1992) and can be less biased than Pearson's correlation coefficient *r* when group sizes are different (McGrath & Meyer, 2006). Cohen's *d* may be interpreted as follows: an effect size

of .30 is considered “small”; an effect size of .50 is considered “moderate”; and an effect size of .80 is considered “large”.

Results

Description of Sample

Chi-Square (χ^2) analyses were conducted to identify any significant differences between the two groups: treatment completers and non-completers (See Table 3.1 for frequencies and chi-square results). There were no differences identified in the majority of variables (see below). The majority of participants were convicted of violent offences, representative of 84.3% of treatment completers and 76.7% of non-completers. No significant differences were identified relating to offence type (see Table 3.2 for frequencies of offence types).

There was a significant difference found in terms of risk of harm, $\chi^2(1, N = 259) = 4.62$, $p = .032$ with 6.5% of treatment completers being identified as high risk in comparison compared to 16.3% of the non-completer group. In this case, Cramer’s $V = .134$ indicating that whilst the association was significant, it represented a weak association (Rea & Parker, 1992). Using odds ratio, the results suggest that an individual who does not complete the programme is 2.81 times more likely to be high risk however this represented a small section of the overall sample.

There was a significant association found between treatment group and Order breach, $\chi^2 (1, N = 259) = 49.67$, $p < .001$. Cramer’s $V = .438$, indicating a relatively strong association (Rea & Parker, 1992). This suggests that, based on odds ratio, a non-completer was 15.02 times more likely breach their Order. This must be viewed with caution, given that individual’s failing to attend the programme leads to programme suspension which may also count as failing to adhere to their Order. This may inherently increase the likelihood of breach.

Notably, 93% of the non-completer group and 88.4% of treatment completers had a history of IPV (convictions, cautions or self-report). Although this difference was not significant, FET (N = 259), p = .590 it is a high frequency of previous IPV. Of the non-completers, 42.5% had previous IPV-related convictions in comparison to 27.9% of treatment completers. However, the differences between the nature of the sample's IPV record (convictions, caution or both or self-report) was also non-significant, χ^2 (3, N = 230) = 6.53, p = .089. There were no significant associations between the type of IPV history in terms of offence categories (contact, non-contact or both), χ^2 (2, N = 229) = 3.70, p = .157.

Reasons for non-completion of the CDVP (N = 43) were varied and included; Poor attendance, n = 24; poor behaviour (disruption/ lack of engagement), n = 6; further offences n = 6; physical or mental health grounds, n = 3; employment issues n = 1; subject to custody due to recall or activation of SSO for reasons unknown, n = 3. Out of the six who had committed further offences, three of these were subject to custody.

Exploratory data analysis

A number of K-S tests found that the majority of psychometric scales were found to be significantly different to a normal population in each group and stage of psychometric testing at a significance level of $p < .01$ (see Appendix eleven, Table i). The data for the RATOS at the pre-treatment in the completer group was D(213) = .066, p = .023. Given that the RATOS variable was found to be normally distributed in all other conditions, a parametric test was utilised. The ECR-R Avoidant scale was found to be normally distributed at all levels. For the list of the 'number of months' variables see Appendix eleven, table ii. Where data was normally distributed on a variable for one group but not another, a non-parametric test was used.

Table 3.1

A table showing the Frequencies (percentages in parentheses) of demographic variables for treatment completers and non-completers.

Variable N(%)	**Treatment completers (N=216)**					**Non-completers (N=43)**					χ^2
Risk of Harm	Medium 202 (93.5)	High 14 (6.5)				Medium 36 (83.7)	High 7 (16.3)				4.62*
Ethnicity	White 169 (78.2)	Asian 8 (3.7)	Black 16 (7.4)	Mixed 9 (4.2)	Other 2 (0.9)	White 39 (90.7)	Asian 0 (0.0)	Black 2 (4.7)	Mixed 2 (4.7)	Other 0 (0)	FET 1.96
Employment	Yes 124 (57.4)	No 82 (38.0)				Yes 13 (30.2)	No 17 (39.5)				3.06
Sentence type	Comm. 149 (69.0)	Susp Sent 60 (27.8)	Licence 7 (3.2)			Comm. 27 (62.8)	Susp Sent 15 (34.9)	Licence 1 (2.3)			.925
Offence category	Contact 128 (59.3)	Non-contact 32 (14.8)	Dual 56 (25.9)			Contact 25 (58.1)	Non-contact 7 (16.3)	Dual 11 (25.6)			.060
Relationship to victim	Ex-partner 52 (24.1)	Current Partner 106 (49.1)	Wife 42 (19.4)	Ex-wife 8 (3.7)	Other 8 (3.7)	Ex-partner 10 (23.3)	Current partner 23 (53.5)	Wife 6 (14.0)	Ex-wife 4 (9.3)	Other 0 (0)	4.75
IPV history	Yes 191 (88.4)	No 25 (11.6)				Yes 40 (93.0)	No 3 (7.0)				FET
IPV Record	Convictions 53 (27.9)	Cautions 50 (26.3)	Both 13 (6.8)	Self-report 74 (38.9)		Convictions 17 (42.5)	Cautions 4 (10.0)	Both 4 (10.0)	Self-report 15 (37.5)		6.53
IPV history offence category	Contact 67 (35.4)	Non-contact 38 (20.1)	Dual 84 (44.4)	N/A 26		Contact 8 (20.0)	Non-contact 9 (22.5)	Dual 23 (57.5)	N/A 3		3.70

Previous violence	Yes: 78 (36.1)	No: 138 (63.9)			Yes 17 (39.5)	No 26 (60.5)			.181
With children	Yes: 133 (61.6)	No: 56 (25.9)	Step-children 27 (12.5)		Yes 31 (72.1)	No 8 (18.6)	Step-children 4 (9.3)		1.71
Child Present at offence	Yes 107 (66.9)	No 53 (33.1)	N/A 56		Yes 27 (77.1)	No 8 (22.9)	N/A 8		1.41
Substance misuse	No 59 (27.3)	Alcohol 128 (59.3)	Drugs 11 (5.1)	Both 18 (8.3)	No 13 (30.2)	Alcohol 23 (53.5)	Drugs 4 (9.3)	Both 3 (7.0)	FET 1.70
History of mental health	Yes 101 (46.8)	No 115 (53.2)			Yes 23 (53.5)	No 20 (46.5)			.651
Marital status	Separated 64 (29.6)	In rel 58 (26.9)	Co-habiting 52 (24.1)	Married 42 (19.4)	Separated 13 (30.2)	In rel 10 (23.3)	Co-habiting 14 (32.6)	Married 6 (14.0)	1.78
Breach	Yes 9 (4.2)	No 207 (95.8)			Yes 17 (39.5)	No 26 (60.5)			49.67**

Note: Not all frequencies equal 100% for each treatment group due to unavailable data. *χ^2 is significant at $p < .05$. ** χ^2 is significant at $p < .00$

Table 3.2

Table showing frequencies of current offence types for treatment completers and non-completers.

Group	Offence Type	Frequency	Percentage (%)
Treatment Completer	Property Damage and Offensive behaviour	6	2.7
	Violent Offence	182	84.3
	Harassment	24	11.1
	Sexual Offence	1	.5
	Other (driving; possession of weapon)	3	1.4
	Total	216	100.0
Non Completer	Property Damage and Offensive behaviour	4	9.3
	Violent Offence	33	76.8
	Harassment	5	11.6
	Sexual Offence	1	2.3
	Total	43	100.0

Due to the non-normality of the majority of variables, the Levene's statistic (F) was obtained based upon the median. Browne and Forsythe (1974) recommend using an adjusted Levene's test on variables that were not normally distributed as this increases the robustness and power of the Levene's test against skewness. Results showed that all variables met the assumption indicating that the variances were equal in the two groups ($p > .05$) (see Appendix eleven, table iii).

The results will now be examined in relation to each hypothesis.

Hypothesis 1

There will be a significant difference in pre-treatment scores for treatment completers and non-completers.

A series of paired tests were conducted to examine differences in pre-treatment means for both groups. Mann Whitney U tests were conducted for the STAXI-2, IRS and ECR-R Anxious scales due to non-normality of the data (see Table 3.3). A Bonferroni correction was applied meaning .003 was used as the alpha level to confirm significance (.05 divided by the number of variables, *N*=19). The Monte Carlo significance levels were observed due to the larger sample sizes within this study (Field, 2009).

Table 3.3

Table showing results of Mann Whitney U tests on pre-treatment psychometric scores showing Means, SD, U *statistic, z score, significance level and effect size (r).*

Variable	**Treatment Group**	***N***	***M (SD)***	***Mdn***	**U**	***z***	***P***	r^1
S-Ang Feeling	Completer	216	51.83 (18.97)	40.00	4250.50	-.908	.354	-.056
	Non-Completer	43	56.14 (22.78)	40.00				
S-Ang Verbal	Completer	216	57.75 (15.05)	50.00	4047.50	-1.76	.082	-.109*
	Non-Completer	43	62.67 (18.51)	50.00				
S-Ang Physical	Completer	216	55.53 (14.35)	50.00	4190.00	-1.63	.111	-.101*
	Non-Completer	43	59.67 (18.01)	50.00				
Total STATE	Completer	216	52.29 (19.27)	40.00	3971.50	-1.54	.121	-.096
	Non-Completer	43	58.26 (22.29)	50.00				
T-Ang Temperament	Completer	216	69.68 (24.52)	75.00	4173.00	-1.06	.290	-.066
	Non-Completer	43	74.26 (22.38)	75.00				
T-Ang Reaction	Completer	216	39.20 (30.52)	30.00	4319.00	-.728	.469	-.045
	Non-Completer	43	43.12 (32.31)	30.00				
Total TRAIT	Completer	216	57.81 (30.76)	60.00	3198.00	-1.62	.109	-.101*
	Non-Completer	43	65.91 (30.18)	70.00				

AX-O	Completer	216	64.84 (29.37)	70.00	4317.00	-.731	.461	-.045
	Non-Completer	43	68.21 (29.17)	70.00				
AX-I	Completer	216	57.67 (29.79)	55.00	4487.50	-.350	.721	-.022
	Non-Completer	43	59.47 (28.20)	60.00				
AC-O	Completer	216	29.02 (27.86)	20.00	4501.50	-.319	.754	-.020
	Non-Completer	43	26.60 (25.68)	20.00				
AC-I	Completer	216	37.36 (25.88)	32.50	4310.50	-.745	.456	-.046
	Non-Completer	43	34.86 (27.28)	30.00				
Anger Expression Index	Completer	216	67.46 (28.52)	75.00	4261.00	-.856	.394	-.053
	Non-Completer	43	70.88 (27.80)	80.00				
IRS Threat	Completer	216	30.07 (7.42)	30.00	4132.50	-1.14	.256	-.071
	Non-Completer	43	28.40 (8.50)	27.00				
IRS Envy	Completer	216	32.83 (7.10)	33.50	4392.50	-.562	.574	-.035
	Non-Completer	43	31.79 (8.30)	34.00				
IRS Dependent	Completer	216	16.47 (6.02)	17.00	4526.50	-.263	.792	-.016
	Non-Completer	43	16.67 (6.37)	19.00				
IRS Sexual Possessive	Completer	216	7.76 (3.58)	8.00	4487.00	-.352	.726	-.022
	Non-Completer	43	8.00 (4.70)	7.00				
IRS Competitive	Completer	216	13.66 (3.27)	14.00	4304.00	-.762	.452	-.047
	Non-Completer	43	13.05 (4.01)	14.00				
IRS Distrust	Completer	216	13.08 (4.10)	13.00	4407.00	-.529	.532	-.033
	Non-Completer	43	12.79 (3.75)	13.00				
ECR-Anxious	Completer	216	3.42 (1.11)	3.44	4350.00	-.656	.513	-.041
	Non-Completer	43	3.54 (.97)	3.67				

Note. Standard Deviations appear in parentheses next to means. [1]The effect size *r* can be approximated using the *z* statistic produced by the Mann-Whitney U test (Field, 2009). The calculation is $z/\sqrt{N}$ and the values for the interpretation of *r* are; .1 small, .3 moderate, .5 large. The minus sign bears no relation to direction of effect. *small association.

There was no statistically significant differences in pre-treatment psychometric scores between treatment completer and non-completers (see Table 3.3). However, small effect sizes were found on STAXI-2 subscales. The results show that non-completers had higher levels of feeling like expressing anger verbally (U = 4047.50, z = -.1.76, p = .082, r = -.109) and physically (U = 4190.00, z = -1.63, p = .111, r = -.101) at the pre-treatment stage when compared to treatment completers. Further, non-completers self-reported having higher levels of overall State anger (experiencing higher intensity of anger in a specific situation) (U = 3971.50, z = -1.54, p = .121, r = -.096) and higher levels of Trait anger indicating they have a disposition to experience angry feelings routinely (U = 3198.00, z = -1.62, p = .109, r = -.101).

Independent samples t-tests were also conducted to examine any significant differences between treatment completers and non-completers at the pre-treatment stage on the ECR-R Avoidant and RATOS variables (see Table 3.4).

Table 3.4

Table showing results of independent t-tests reporting significance level and effect size (Cohen's d).

Variable	**Treatment Group**	***N***	***M (SD)***	***T***	***df***	***p***	***d***
ECR-R Avoidant	Completer	216	3.04 (0.92)	.315	257	.002*	.537**
	Non-completer	43	3.54 (1.01)				
RATOS	Completer	213	42.38 (10.65)	.534	254	.594	.091
	Non-completer	43	43.33 (9.91)				

Note. *Significant at < .01. ** Cohen's *d* (Cohen, 1988, 1992) may be interpreted as follows: an effect size of .30 is considered "small"; an effect size of .50 is considered "moderate"; and an effect size of .80 is considered "large."

A Bonferroni correction was applied (.05 divided by the number of variables = 2) yielding an adjusted *p* value of .025 to confirm significance. The results indicate that non-completers had higher levels of insecure avoidant attachment, i.e. experience stronger feelings of being uncomfortable in closeness to others, opening up to others and depending on others, $t(257) = .32$, $p = .002$, $d = .52$. This demonstrates a moderate effect size.

There were no significant differences found between treatment completers and non-completers in their self-reported presence of denial, minimisation, victim-blaming and failure to take responsibility ($t(254) = .53$, $p = .59$, $d = .09$). Cohen's *d* was calculated with the *SD* appropriately weighted where sample sizes were not equal (Zakzanis, 2001).

The pre-treatment scores on the RATOS for both the treatment completer and non-completer group were also compared to the normed CDVP data (NOMS Test Battery, 2012) using independent t-tests. The results found no significant differences between the treatment group, $t(745) = 1.06$, $p = .29$, although yielded a small effect size of $d = .17$. Similarly, the non-completer group did not significantly differ from the normed population, $t(915) = 1.20$, $p = .23$ yet yielded a small effect size of $d = .10$.

Whilst differences were not observed on all variables, differences were identified by the effect size and one significant result at $p < .025$ therefore H1 was accepted.

Hypothesis 2

There will be a significant difference in attitude changes among treatment completers (pre, post and six-month post treatment).

A series of Friedman's ANOVA and one-way related samples ANOVA were conducted to examine differences in scores at the pre-treatment, post-treatment and six-month post-treatment stage amongst treatment completers dependent upon normality of the data. It is of note that *N* varies throughout due to the level of attrition at the six-month post-treatment

stage (N =50 out of the original 216 participants). Whilst this sample size may be considered large (Field, 2009) the results cannot be generalised to all treatment completers within this sample.

A Bonferroni adjustment was made based upon the number of variables examined using the Friedman ANOVA (19) and the number of comparisons made at each post-hoc stage (3). This yielded a new alpha value of p = .001 (.05/19 = .003. .003/3 = .001). Where a significant main effect was found, further post-hoc tests were carried out using Wilcoxon signed-rank tests (Wilcoxon, 1945) and related sample t-tests where appropriate. Post-hoc tests were conducted on all participants to minimise loss of the quantity of the data collected. Therefore N is larger within pre-treatment to post-treatment comparisons (N = 216).

The Friedman ANOVA found significant main treatment effects of 11 out of the 19 non-normal variables; S-Ang/F $\chi^2(2, N = 50) = 17.11$, $p < .001$, S-Ang $\chi^2 (2, N= 50) = 17.96$, $p < .001$, T-Ang/T $\chi^2(2, N = 50) = 14.22$, $p = .001$, T-Ang, $\chi^2(2, N = 50) = 18.30$, $p < .001$, AX-O $\chi^2(2, N = 50) = 14.60$, $p = .001$, AC-O $\chi^2(2, N = 50) = 19.97$, $p < .001$, AC-I $\chi^2(2, N = 50) = 23.75$, $p < .001$, Anger Expression Index $\chi^2(2, N = 50) = 23.42$, $p < .001$, IRS Threat to exclusivity $\chi^2(2, N = 49) = 21.32$, $p < .001$, IRS Dependency $\chi^2(2, N = 49) = 13.29$, $p = .001$ and IRS Competitiveness/Vindictiveness $\chi^2(2, N = 49) = 13.66$, $p = .001$ (see Appendix twelve for full statistical outputs).

No significant effects were found on S-Ang/V, S-Ang/P, T-Ang/R, AX-I, ECR-R Anxious, IRS Sexual Possessiveness, IRS Envy or IRS Distrust therefore no post-hoc comparisons were made (see Appendix twelve).

Effect sizes were not calculated for the main analyses due to the difficulty in converting a chi-square statistic that has more than one degree of freedom (Field, 2009) therefore effect size r was calculated on post-hoc comparisons to inform relationships.

Wilcoxon signed-rank post hoc analysis revealed a number of significant effects amongst the 11 variables above (See Table 3.5).

As can be seen from Table 3.5, the majority of the comparisons revealed a significant finding and moderate to large effect sizes. Note: The T statistic denotes the smaller of the two sums of ranks for each test.

Table 3.5

Table showing post-hoc tests using Wilcoxon signed-rank test amongst non-normal variables including T statistic, the value of Z, its significance level and effect size r.

Variable	**Treatment stage comparison**	***N***	***Ties***	***T***	***z***	***p***	***r***
S-Ang/F	Post – Pre	216	115	572.00	-6.81	.001**	-.463
	Follow up - Pre	50	29	12.50	-3.59	.001**	**-.508**
	Follow up - Post	50	39	22.50	-.940	.347	-.133
S-Ang	Post – Pre	216	107	879.00	-6.43	.001**	-.437
	Follow up - Pre	50	24	41.50	-3.42	.001*	-.484
	Follow up - Post	50	37	36.00	-.668	.504	-.094
T-Ang/T	Post – Pre	216	54	4136.00	-4.13	.001**	-.281
	Follow up - Pre	50	15	93.00	-3.64	.001**	**-.515**
	Follow up - Post	50	21	133.00	-1.34	.066	-.190
T-Ang	Post – Pre	216	31	5687.00	-4.00	.001**	-.272
	Follow up - Pre	50	6	130.00	-4.26	.001**	**-.603**
	Follow up - Post	50	13	213.50	-2.09	.037	-.296
AX-O	Post – Pre	216	27	5651.00	-4.42	.001**	-.301
	Follow up – Pre	50	3	178.00	-4.09	.001**	**-.579**
	Follow up - Post	50	10	275.00	-1.82	.069	-.257
AC-O	Post – Pre	216	29	3548.50	-7.08	.001**	-.482

	Follow up – Pre	50	1	193.00	-4.18	.001**	**-.591**
	Follow up - Post	50	11	316.00	-1.03	.301	-.146
AC-I	Post – Pre	216	19	2907.00	-8.55	.001**	**-.582**
	Follow up – Pre	50	0	150.00	-4.71	.001**	**-.667**
	Follow up - Post	50	11	181.00	-2.93	.003*	-.414
AX Index	Post – Pre	216	22	3608.00	-7.47	.001**	**-.508**
	Follow up – Pre	50	4	123.50	-4.56	.001**	**-.645**
	Follow up - Post	50	11	225.00	-2.31	.021	-.327
IRS Threat	Post – Pre	216	19	6647.50	-3.88	.001**	-.264
	Follow up – Pre	49	6	149.00	-3.92	.001**	**-.560**
	Follow up - Post	49	5	279.50	-2.52	.012	-.360
IRS Dependent	Post – Pre	216	41	4627.00	-4.59	.001**	-.312
	Follow up – Pre	49	9	197.00	-2.87	.004	-.410
	Follow up - Post	49	10	233.50	-2.19	.028	-.313
IRS Competitive	Post – Pre	216	41	5848.50	-2.76	.006	-.188
	Follow up – Pre	49	10	184.50	-2.89	.004	-.413
	Follow up - Post	49	13	271.00	-.993	.321	-.142

Note: r; +/- .1 = small, .3 = moderate, .5 = large. The minus sign bears no relation to direction of effect for r. Figures in bold represent large effect sizes. *significant at ≤ .003 ** significant at $p < .001$.

The S-Ang/F domain yielded 115 tied scores when comparing pre-treatment to post-treatment. This reduces the power of the test as the results are based upon 101 of the original sample. This was also observed in post-treatment to six-month post treatment and pre-treatment to follow up stages (29 ties and 39 ties respectively of $n = 50$). This domain found the highest number of ties.

In relation to IRS Threat, higher scores on the IRS scale indicated lower levels of romantic jealousy, a significant increase in scores at the post-treatment stage indicated lower

levels of reported jealousy whereas a reduction in scores would indicate an increase in jealousy levels.

One-way repeated-measures ANOVA's were applied to two of the scales; ECR-R Avoidant (N = 50) and RATOS (N = 47). On the ECR-R Avoidant scale, Mauchly's test indicated that the assumption of sphericity had been violated ($\chi^2(2, N = 50) = 6.27$, $p = .044$) therefore Greenhouse-Geisser correction was applied ($\varepsilon = .891$). Field (2012) highlighted that whilst the violation of the sphericity assumption can reduce the power of the ANOVA, the Greenhouse-Geisser correction leads to a more conservative F statistic.

No significant effects were found when examining at the pre-treatment (M = 2.99, SD = .91), post-treatment (M = 2.88, SD = .85) and six-month post treatment (M = 2.79, SD = .84) stages, $F(1.78, 87.31) = 1.61$, $p = .209$, $\eta_p^2 = .032$. The effect size of partial eta squared indicates that ECR-R avoidance accounted for 3.20% of the variance within the sample and its associated error.

The assumption of sphericity was also violated on the RATOS ($\chi^2(2, N = 47) = 7.43$, $p = .024$) and the Greenhouse-Geisser correction was applied ($\varepsilon = .868$). Significant differences were in an individual's self-reported levels of minimisation, denial and blame as measured by the RATOS at each treatment stage ($F(1.74, 79.85) = 34.95$, $p < .001$) with a moderate effect size, $\eta_p^2 = .432$.

Post-hoc pairwise comparisons with a Bonferroni adjustment found significant differences at $p < .001$ level between pre-treatment RATOS scores (M = 42.28, SD = 11.01) and post-treatment scores (M = 33.87, SD = 10.87) and between pre-treatment and six-month post treatment scores (M = 32.36, SD = 12.48) but there were no significant differences between post-treatment and six-month post-treatment scores ($p = .481$).

Given the statistical differences found between the three treatment stages amongst several psychometric scales, H2 was accepted.

Hypothesis 3

There will be significant differences between treatment completers and non-completers in relation to re-offending and reconviction rates.

The mean follow-up period for treatment completers was 8.49 months (*SD* = 4.61 months, *Mdn* = 8.00 months). The mean follow-up period for non-completers was 9.91 months (*SD* = 5.61 months, *Mdn* = 10.00 months). Note that the *Mdn* is represented as this data was found to be non-normally distributed, $D(216) = .098$, $p = .012$ and $D(43) = .294$, $p < .001$ respectively. There was no significant difference between this length of time among the two groups, $U = 3910.50$, $z = -1.64$, $p = .101$, $r = -.10$.

When examining the total sample, 23.9% (62 out of $N = 259$) of participants were convicted of an IPV-related offence (8.1%) or were involved in an 'alleged IPV re-offence' (15.8%). Out of the 216 treatment completers, 26% ($n = 56$) were identified as recidivists in some form (either general or IPV related offending) compared to 39.6% of non-completers ($n = 17$), $\chi^2 (3, N = 259) = 16.59$, $p = .001$. Cramer's $V = .253$ indicated a moderate association.

There was a significant association found between treatment condition (treatment completers versus non-completers) and non-IPV related reconviction, $\chi^2 (1, N = 259)$, 6.77, $p = .01$ although Cramer's $V = .16$ indicated a weak association. To adjust for the Chi-square assumption violation noted above, the Fisher's Exact Test (FET) significance level was also observed, $p = .021$. In terms of percentage frequencies, 20.9% of non-completers were convicted of a further general offence (not IPV related) compared to 7.9% of treatment completers.

Results demonstrated that 20.9% of non-completers were reconvicted of an IPV offence compared to 5.6% of treatment completers which was a statistically significant difference, $\chi^2 (1, N = 259) = 11.38$, $p = .001$. Because 25% of cells (1 cell) had expected frequencies of less than five, the FET was observed and showed significance at $p = .003$ with

Cramer's $V = .21$ demonstrating a moderate association. This indicates that, based on odds ratio, non-completers may be 4.51 times more likely to be reconvicted of an IPV offence than treatment completers. The mean number of months until IPV reconviction between the two groups (treatment completer group, $M = 6.75$, $SD = 4.18$; non-completer group, $M = 10.44$, $SD = 7.59$) did not significantly differ, $t(19) = 1.43$, $p = .169$, $d = .662$. However, Cohen's d indicated a moderate effect size.

Amongst the treatment completers, 16.2% were identified allegedly re-offending in an IPV context compared to 14% of non-completers. This percentage was non-significant, χ^2 (1, $N = 259$) = .14, $p = .71$. In addition, there were no significant differences found between the sources of information used to ascertain alleged re-offending (DAU, Police or self-report) between the two groups, FET ($N = 41$), $p = 1.00$. The FET was used as 50% of cells contained expected frequencies of less than five due to the small sample size. Further, there was no significant difference in whether DAU information was available for either group, χ^2 (1, $N = 259$) = 1.19, $p = .28$ (see Appendix thirteen for table of all recidivist and non-recidivist frequencies and Chi-Square results).

Further IPV reconviction and alleged IPV re-offending was carried out against the same victim in the majority of cases for both treatment completer (See Table 3.6).

Table 3.6

Table showing percentages of recidivists who offended (alleged or actual) against the same victim as original offence.

Treatment Group	IPV Reconviction		Alleged IPV re-offending	
	n	Offended against same victim	*n*	Offended against same victim
Treatment completer	12	91.7%	35	80.0%
Non-completer	9	77.8%	6	83.3%

Examining Overall IPV Recidivism using Survival Analysis

Due to the low figures, IPV reconviction and alleged IPV re-offending rates were combined to form 'Overall IPV recidivism'. A K-S test on the number of months until 'overall IPV recidivism' revealed a normal distribution (treatment completers, $D(45) = .116$, $p = .156$; non-completers, $D(11) = .175$, $p = .200$). However, the variable was revealed to violate the assumption for homogeneity of variance, $F(1, 54) = 6.22$, $p = .026$. Therefore a Mann-Whitney U was conducted and no significant differences were found between treatment completers and non-completers and number of months until 'overall IPV recidivism', $U = 221.00$, $z = -.548$, $p = .598$, $r = -.073$.

There were no significant difference between the two groups and overall IPV recidivism, χ^2 (1, $N = 259$) = .477, $p = .490$ (25.6% non-completer recidivism versus 20.8% treatment completer recidivism).

Survival analysis is a statistical procedure that involves the modelling of time or distance to a given event (Garson, 2012) providing more insight. In this case, the survival function *S(t)* is modelled on time in months to a participants' overall IPV recidivism using the Life Table. The data was analysed and plotted (Figure 2).

The median survival rate (the point at which 50% of the sample survive) for treatment completers ($n = 45$) was 6.38 months in comparison to 7.50 months for non-completers ($n = 11$). Intriguingly, Figure 2 indicates that treatment completers had a lower survival rate in comparison to the non-completer group. However, a pairwise comparison using the Wilcoxon Gehan statistic found that these differences were non-significant, $T(1) = .300$, $p = .584$. The overall mean survival time amongst both groups was 7.04 months ($SE = .670$; Median = 6.00).

The Survival Analysis output table indicated that more participants re-offended (actual or alleged) within the first six months of treatment completion as presented in Table

3.7. Further exploration of survival probabilities (and *SE*'s) for treatment completers is presented in Table 3.8. This data was not presented for those who recidivated in the non-completer group due to lack of meaningfulness as a result of small sample size.

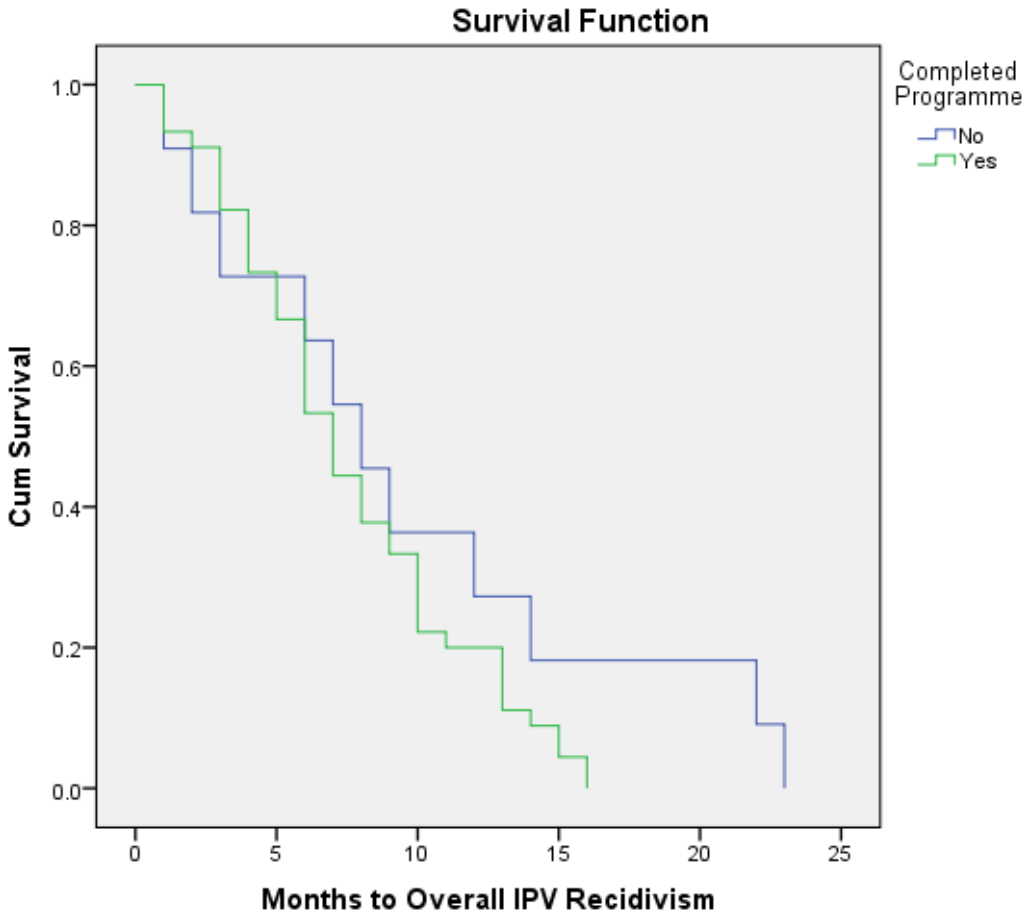

Figure 2. Plot of the proportions of treatment completers and non-completers who reoffend (alleged IPV re-offending and IPV reconviction combined) at varying times.

Table 3.7

Table showing number of treatment completers re-offending within set time frames.

Group	Time to Overall IPV Recidivism event		
	0 – 6 months	7 – 12 months	13 – 15 months
Number of Treatment Completers	25	15	5

Table 3.8

Table showing the cumulative proportions of survival at the end of each interval amongst treatment completer recidivists.

Group	Time Interval				
	< 1 month	3 months	6 months	9 months	12 months
Treatment Completer Recidivists	93% (4%)	73% (7%)	44% (7%)	22% (6%)	11% (5%)

Note. Standard Error in parentheses.

Table 3.8 indicates that out of all recidivists, 73% do not go on to reoffend (alleged or actual) at the three month stage six after programme completion (± Standard Error). Conversely it shows that out of all treatment completer who recidivated (n = 45 out of 216), between 72% and 84% of those had recidivated nine months after the programme completion.

The results for H3 suggest that there was a significant difference in the frequency of non-IPV reconviction although this was not a strong association. There was a significant moderate association between IPV reconviction and whether an individual had completed the programme. This finding disappeared when alleged IPV re-offending was observed and when reconviction and re-offending was combined. H3 was partially accepted.

Hypothesis 4

There will be a significant difference among recidivists and non-recidivists in relation to their demographics and post-psychometric scores (for treatment completers).

A number of exploratory analyses on the variable of 'Overall IPV recidivism' amongst treatment completers' post-treatment psychometric data as these were considered to be new groups separate from the 'treatment completer' and 'non-completer' distinction made in the previous hypotheses. Post-treatment scores were examined as opposed to six-month post-treatment scores due to the level of attrition enabling data to be analysed for all recidivists and non-recidivists in the treatment-completer group, $n = 171$ and $n = 45$ respectively. A table of all statistical outputs for examining assumptions on the data for IPV Overall Recidivism can be found in Appendix fourteen, Tables i and ii**.** It is important to note the unequal sample sizes throughout the analyses. Where variables were normally distributed, independent t-tests were conducted. Although the sample sizes are unequal, Field (2009) highlighted that t-tests are robust to this where there is homogeneity of variance which was case for the ECR-R Avoidant and RATOS scales, $F(1, 214) = 2.40$, $p = .123$ and $F(1,211) = .281$, $p = .597$ respectively.

No significant differences were found between the means for Overall IPV recidivists on the ECR Avoidant variable ($M = 2.68$, $SD = .737$, SE = .110) and non- recidivists ($M = 2.92$, $SD = .890$, SE = .068), $t(214) = 1.67$, $p = .096$; however a small effect size was observed, $d = .281$. There were no other significant differences found between the means on post-treatment RATOS scores for recidivists ($M = 32.00$, $SD = 11.98$, $SE = 1.81$) and non-recidivists ($M = 34.14$, $SD = 11.72$, $SE = .901$), $t(211) = 1.08$, $p = .284$, with an effect size of $d = .183$.

Where variables were significantly non-normal, Mann Whitney U tests were employed with a Bonferroni adjusted *p* value (.003 as in the above hypotheses) which does not require equal sample sizes (Field, 2009).

In relation to the IRS scales, all variables were found to be non-significantly different between groups however in the IRS Threat to Exclusivity domain, results indicated a small effect size where recidivists (*Mdn* = 31.00) reported higher levels of jealousy in relation to this scale in comparison to non-recidivists (*Mdn* = 33.00), $U = 3210.00$, $z = -1.71$, $p = .086$, $r = -0.116$. All other non-normal variables were found not to significantly differ (see Appendix fifteen for a table displaying all Mann Witney U outputs).

Recidivists (*Mdn* = 31.00) years were found to be younger than non-recidivists (*Mdn* = 36.00 years), $U = 3258.50$, $z = -1.58$, $p = .111$ with a small effect size of $r = -.107$.

Chi-Square tests were conducted to explore differences on demographic variables amongst treatment completer recidivists and non-recidivists (see Appendix sixteen). Marital status amongst treatment completer recidivists was significantly different to non-recidivists at the time of their original offence, $\chi^2 (3, N = 216) = 10.34$, $p = .016$, with 44.4% of recidivists in a non-cohabiting relationship with the victim in comparison to 22.2% of non-recidivists. Further, 31.6% of non-recidivists were separated from the victim in comparison to 22.2% recidivists. Cramer's $V = .219$ showed a moderate association.

H4 was accepted and implications of findings will be discussed.

Discussion

The aim of this study was to expand on current literature relating to the effectiveness of IPV interventions for male perpetrators. This was addressed through examining a probation sample of IPV treatment completers' psychometric measures and recidivism rates. Treatment completers were compared to non-completers on a number of attitudinal, behavioural and demographic variables. Differences between treatment completer recidivists and non-recidivists were investigated with an aim of gaining further insight into the characteristics associated with re-offending. Whilst usual to discuss the hypotheses in logical order, it was considered more useful in this case to highlight the themes of the current study brought out by the hypotheses. Consequently, the summary and evaluation of the main findings will first discuss the differences between groups incorporating H1 and H4, following on to consider the effect of treatment upon psychological variables incorporating H2 and finally, examining recidivism rates as informed by H3.

Summary and evaluation of main findings

Differences between groups

In relation to H1, there were no significant differences pertaining to demographic variables when comparing treatment completers and non-completers at the pre-treatment stage. However a significantly higher number of non-completers were categorised as high risk, supporting McMurran and Theodosi (2007). Higher risk individuals may present as more challenging which is consistent with the RNR model (Andrews et al., 1990; Andrews & Bonta, 2010a). This is again supported by significantly more non-completers breaching their orders which may reflect a lack of motivation (Bowen & Gilchrist, 2006). Additionally, non-completers self-reported higher levels of feeling like expressing their anger verbally and physically, and overall Trait anger. Such individuals may be more likely to experience anger

in situations and be engaging in outward verbal aggression. This could result in less positive engagement and ultimately programme suspension (McMurran & Theodosi, 2007).

A moderate effect was found with regards to non-completers reporting an increased avoidant attachment style. This could potentially act as a mediating factor between an individual, the attachment style and their level of comfort and willingness to open up and settle within a group. The role of group cohesiveness has been identified an important factor in positive outcomes in IPV intervention (Bowen, 2010; Taft et al., 2003). Unexpectedly, recidivists reported slightly lower levels of avoidant attachment. Interpersonal dependency, which may be considered as more closely linked to the construct of anxious attachment, has been associated with post-programme recidivism amongst IPV offenders (Bowen et al., 2005). Given these findings and previous literature highlighting the complexity of attachment and IPV (Doumas et al., 2008) the field would benefit from further research regarding its role in both IPV offending and treatment completion and engagement.

Few differences were observed when exploring demographic variables amongst treatment completer recidivists versus non-recidivists as examined in H4. However, recidivists tended to be slightly younger in age, with a small effect size, which is consistent with longitudinal studies that have suggested IPV may decrease with age (Shortt et al., 2012). This could also be considered in relation to literature relating to the age-crime curve (Farrington, 1986; Loeber & Farrington, 2014). The finding of an age-crime curve has been consistently found across different countries whereby offending behaviour starts in late childhood, peaks between aged 15-19 years and declines steeply to the early 20's and more steadily thereafter (Farrington, 1986). The current sample depicts an older age than the age-crime curve would suggest (31 and 36 years old for recidivist- and non-recidivist treatment completers respectively). The reality is more complex and there is research evidence that IPV

does not follow the same age-related trends (Wilke & Vinton, 2005). This highlights a need for further research in understanding direct and indirect relationships between age and IPV, particularly older-aged victims and perpetrators (see Chapter five). Indeed, cessation of offending is far more complex than age alone and must considered within a context of other psychosocial, attitudinal and behavioural factors (Loeber et al., 2012; Walker et al., 2013).

Post-treatment psychometric comparisons found that treatment completer recidivists self-reported higher levels of perceived threat to exclusivity of a relationship than those who did not. This suggests that individuals who go on to reoffend may experience higher levels of jealousy of a (ex) partner. Despite a significant reduction in reported levels of jealousy when examined at each treatment stage, it is possible that there is a sub-set of individuals who did not achieve such change within this domain. Conclusions cannot be made based on the current study's methodology. There is a need for further exploration into attitudinal and behavioural change at an individual level (Bowen et al., 2008). This finding demonstrates the importance of effectively identifying criminogenic needs and assessing progress throughout a programme.

When considering group differences, more confident conclusions were made when examining treatment completers and non-completers as opposed to treatment completer recidivists versus treatment completer non-recidivists.

Effectiveness of CDVP on attitudinal change for treatment completers

The results of H2 were obtained through examining pre-, post- and six-month post-treatment questionnaires. The findings demonstrated a number of moderate and large effect sizes between treatment stages. Large effect sizes were observed relating to STAXI-2 scales and subscales indicating that participant's reported; an increase in their ability to control their

anger outwards and inwards; a reduction in feeling like expressing anger verbally and a reduction in experiencing anger routinely (Trait anger). This finding is positive given that higher levels of Trait anger and difficulties with anger expression has been observed among IPV offenders (Barbour et al., 1998). There were a large number of ties seen in the State anger scale showing no reported change following treatment. However, this scale is reflective of an individual's present state and is influenced by a range of situational factors. It could be argued that anger expression/control and trait anger subscales may be more relevant to evaluating psychological change (see Chapter four).

Whilst a significant improvement in individuals' level of responsibility taking was found (as identified by the RATOS), the sample of treatment completers did not highlight concerns relating to denial, minimisation or blame at any treatment stage. Mean samples scores were below the psychometric guidance of 75 which are indicative of firm views and indeed below 50, indicating some presence of a lack of a responsibility taking. This raises the question as to the relevance of responsibility taking being targeted as a criminogenic need amongst this specific sample of IPV offenders despite improvements being observed.

The significant reduction in Treatment completers perceived threat to the exclusivity of a relationship, dependency, competitiveness and vindictiveness is promising given evidence that jealousy is an identified risk factor for IPV perpetration (Fernandez-Montalvo et al., 2005; Puente & Cohen, 2003). Reported levels of sexual jealousy at the pre-treatment stage were much lower than on other scales within the IRS, indicating higher levels of reported jealousy in this domain. There was no significant effect found within this specific domain highlighting further need to explore methods of addressing sexual jealousy within IPV treatment given evidence that it may impact the severity of IPV (Medeiros & Straus, 2006b). Furthermore, the IRS does not have norms or markers to indicate pathological levels

of jealousy which would be of clinical value when considering treatment needs for an IPV offender.

With regards to attachment style, it was evident that there was a reported improvement towards secure attachment on both the anxious and avoidant scales. However, neither scale reached statistical significance. The lack of treatment effect suggests that this area was not successfully targeted by the CDVP which has implications given that non-completers had significantly higher levels of a self-reported avoidant attachment style at the pre-treatment stage.

The above is largely consistent with Stewart et al., (2005) and Connors et al. (2012) although direct comparisons cannot be made due to sample and procedural differences. The results relating to attitudinal change indicate a positive treatment effect.

Effectiveness of CDVP on recidivism

Whilst a number of studies have found reported psychological characteristic change following programme completion (Bowen et al., 2008) it is important to focus upon multiple outcome measures in order to enhance the robustness of claiming an effective IPV intervention.

The results pertaining to H3 indicate the rate of IPV recidivism (actual or alleged) as just under a quarter of the whole sample in the current study which is fairly consistent with the literature (Rosenfeld, 1992) and slightly less compared to Stover et al. (2009). Further, non-completer's recidivated in some form (IPV reconviction, IPV re-offending, general reconviction) at a significantly higher frequency than treatment completers, yielding a moderate effect. Out of the whole sample, 8.1% were reconvicted of a further IPV offence within the time of follow-up which reflects similar rates to that of Rosenfeld (1992) when using official reports. Whilst non-completers were more likely to be reconvicted compared to

treatment completers, this finding reflected just nine and 12 participants from the original sample size respectively. It is therefore deemed inappropriate to make solid inferences from the effect size result as the low numbers reduce power and generalisability. The aggregation of IPV reconviction and alleged re-offending will be discussed below. The findings do however reflect similarly to Stewart et al.'s (2005) study.

Although significant differences were found between treatment groups when considering IPV reconviction, this became non-significant when IPV re-offending was examined. This is consistent with other studies and reviews (Feder & Wilson, 2005). Whilst a combination of official reconviction figures and alleged IPV incident figures were utilised, this could lead to over- or under-representation of IPV re-offending. As discussed in Bowen et al. (2008) the inclusion of 'alleged IPV re-offending' could have been impacted by a number of biases and can only be considered as a proxy for actual behaviour. For example, a disclosure of a further IPV incident by a perpetrator to an OM may be influenced by level of trust or the extent of the individual's awareness of the behaviour as abusive. Furthermore, subsequent action by an OM could be influenced by experience, perceived severity or compliance with an order. Information detailing who contacted the Police in such cases was not available and findings must be considered within the context of the potential biases a witness to an alleged IPV incident may have (Bowen et al., 2008).

Due to the low base rate of IPV re-offending and reconviction in this sample, these were combined and examined. This did not yield significant differences in relation to treatment completion and rates of overall IPV recidivism. Indeed, the overall IPV recidivism rates were lower than in comparison to other treatment effectiveness literature (Stover et al., 2009). The aggregation of IPV reconviction and IPV alleged re-offending eliminated the ability for data to be analysed regarding the context and severity of these further incidents.

Further research into factors associated with recidivism would benefit from exploring this given the complexities of risk factors associated with IPV perpetration, severity and its implications for desistance (Roberts et al., 2010; Theobald & Farrington, 2012; Walker et al., 2013).

There were no differences in the source of information between the treatment completer group and non-completer group. However, it was apparent that more non-completers were reconvicted of an IPV offence despite treatment completers reporting similar levels of alleged IPV re-offending when compared to non-completers. For example, 16.2% of treatment completers reported alleged IPV or were involved in police callouts whilst a further 5.6% received convictions. This was in comparison to 14% of non-completers reporting further alleged IPV incidents and 20.9% being reconvicted for a further IPV related offence. It could be considered that those who have completed the programme have more interaction with their OM's and engage in IPV specific post-group work which could provide more opportunity for disclosure. Another explanation may be that non-completers may be less compliant leading to more direct involvement with the police as opposed to approaching their OM to disclose difficulty and seek support. This is supported in the finding of more non-completers breaching the conditions of their Orders.. A similarly unexpected result that could also be explained by the above was that treatment completers were found to re-offend at a slightly faster rate than non-completers when examining the Survival Analysis although this difference was non-significant. The Survival Analysis findings also descriptively highlight that more 'overall IPV recidivism' among treatment completers occurs within the first six months of completing the programme which reflective of Stover et al.'s (2009) results.

Limitations of the study

Although this study has provided insight into attitudinal and behavioural outcomes in IPV treatment, the findings must be considered in light of several sampling and methodological limitations (Archer, 2000). Whilst some may be viewed as inherent within the field, others are related to this particular study.

Firstly, the retrospective nature of the study limits the ability to make causal predictions or determine causal relationships. It is suggested that a prospective design is considered and employed before any firm conclusions can be made.

The study utilised a quasi-experimental design as opposed to the random assignment of cases to a treatment or control condition. This reduced the internal validity of the study. Namely, there is reduced confidence that observed differences were solely as a result of the intervention (Feder et al., 2011).

Employing a non-completer versus treatment completer design may increase the likelihood of confounds as drop-out is likely to be attributable to other important variables (Feder & Wilson, 2005). In an attempt to reduce this, a number of variables were examined, including substance misuse and mental health which has been found to be associated with increased drop-out (Bowen & Gilchrist, 2004). Only risk level and breaching of an Order distinguished completers and non-completers (McMurran & Theodosi, 2007). It is appropriate to note that treatment drop-out can also be linked to protective factors such as gaining employment (Ward & Maruna, 2007). Other reasons may include disruptive behaviour or being recalled to prison which may be indicative of a higher risk profile (Olver et al., 2011). The varying reasons for drop-out indeed indicate heterogeneity amongst this sample group and such differences should be taken into consideration when making conclusions regarding the results of the current study.

Experimental methods are not without critique in the areas of feasibility and ethics, given the potential risk of harm to the victim if a perpetrator was withheld treatment (Devine, Wright, & Joyner, 1994). Conversely, Feder et al. (2011) argue that this concern assumes that the assigned intervention has been proven to be beneficial, which may not be case. As highlighted, the non-completer participants were not a homogenous group. Employing an 'intention to treat' model within the current study through comparing those who receive treatment (including those who dropped out) to a waiting list control group would have reduced known and unknown group differences (Gupta, 2001). Such an approach would result in a reduction in Type 1 error due to more conservative estimates (Gupta, 2001). Further, an examination of dosage of treatment in the treatment groups may be viewed as a more comprehensive approach, reducing confounding variables linked to drop-out (Feder et al., 2011).

A large and representative sample increases the robustness of a study (Archer, 2000). The group sizes were unequal regarding treatment completers and non-completers although statistical tests were considered robust due to the large sample size utilised in this study (Field, 2009). The small sample of treatment completer recidivists and non-recidivists mean that conclusions relating to re-offending cannot be considered further than the current study. Furthermore, the sample consisted predominantly of white males which limit the generalisability to the culturally diverse population at large. The prevalence of IPV has been found to cross all ethnicities and research samples benefit from reflecting this. A generalisation across locations cannot be made as the current study focussed upon one area. Whilst CDVP is accredited which sets criteria and procedures of delivery, the standard of delivery may differ in each probation area.

There could be a range of possible confounds to the significant finding relating to IPV reconviction including; variations in follow-up time, relapse into substance misuse,

mental health difficulties and life stressors. This study did not have the scope to investigate psychological and situational factors relating to any further offending and therefore associations cannot be solely linked to the effect of treatment completion. Ideally, the groups would have been matched on several variables to reduce for potential confounds however the small sample size did not allow for this.

The majority of the psychometric measures used within the current study possessed a good level of validity and reliability although further research should be conducted on the RATOS. A full critique of the STAXI-2 is provided in Chapter four highlighting the strengths, limitations and implications for results relating to the measure in this study. This highlighted the continued questions concerning the relationship between assessed anger, continued aggressive behaviour and the complexity of different types of IPV (Murphy et al., 2007). Arguably, the STAXI-2 may not adequately capture the experiences and expression of anger in a manner that is applicable to all IPV offending (see Chapter four).

Relying solely on self-report measures to evaluate attitudinal change may introduce bias and inaccuracies particularly within offender populations (Foley et al., 2002; James, 2005; see Chapter four for further discussion of such implications). There is no psychometric measure to assess socially desirable responding in the CDVP psychometric battery which can reduce the internal validity of the study. That said; an individual's self-report does give insight into that individual. Ideally, self-report questionnaires should be used in conjunction with others methods such as observations and documentation reviews. The use of partner reports within this study could have provided a more objective perspective to this current study (Tjarden & Thoennes, 2000) however this methodology isn't without its' biases. Unfortunately, this information was unavailable due to issues surrounding confidentiality.

The findings relating to recidivism must be considered in light of the difficulties that are well-documented more generally and specifically within the field of IPV (Saunders, 2008;

Tjarden & Thoennes, 2000). IPV can take many forms (see Chapter one) and the identification of ongoing intimidation or controlling behaviours within an intimate relationship may be more problematic to uncover (Johnson, 2006). This is increasingly difficult without victim reports. It is important to note that mutual violence may be occurring for those in treatment (or dropping out) (Dixon & Graham-Kevan, 2011) and this has implications in evaluating the context of recidivism.

The method of data collection and data access for this study meant that follow-up periods varied for participants. Whilst follow-up for some participants was up to 28 months, it was much shorter for others. Most follow up periods are a minimum of 12 months (Babcock et al., 2004; Bowen et al., 2008). Those who were lost during the follow-up consisted of individuals whose order's had expired. An individual was therefore no longer traceable unless they had been reconvicted of a further offence.

Practical application and future directions

Despite the observed limitations, the findings obtained in this paper highlight some implications for future research, risk management and intervention in the field of IPV.

Findings relating to the role of avoidant attachment highlight the potential benefit of facilitators to receive attachment-theory training as a means of understanding insecure attachment within the context of the function of offending and general interpersonal styles. Further, this may encourage facilitators to work more responsively to those who may be less forthcoming with their emotions or difficulties which will be particularly useful to apply to in the facilitation of Modules two, three and four (awareness; emotional management; social skills) in CDVP. In addition, facilitators should continue to be mindful of the dyadic nature of IPV and the possibility for mutual violence given evidence for gender symmetry (see Chapter one). Within Module two of CDVP, group members present an autobiography as a means of

increasing awareness and insight into the origins and pattern of their abusive behaviours. The recognition of being both a perpetrator and a victim will encourage more open and empathic questioning and the programme material to be delivered in a manner that is reflective of an individual's circumstances.

The high number of IPV perpetrators who were under the influence of alcohol and experienced mental health difficulties reinforce the requirement for IPV treatment target a wider context of an individual's functioning (Stover et al., 2009). Facilitators' awareness of this will increase the holistic nature of IPV treatment and reduce the more ideological approach that has received criticism (Bowen, 2011).

This study demonstrated that at least 134 children were present during IPV offences. This is a large and concerning figure given that witnessing family violence has been identified as a risk factor in the continuation of abuse (Dixon, Browne & Hamilton-Giachritsis, 2005; Roberts et al., 2010). All professionals involved in the assessment, treatment and management of an IPV perpetrator should therefore consider the co-occurrence of IPV and child maltreatment as highlighted by Dixon and Graham-Kevan (2011). This includes understanding an individual's use of IPV within the familial context through close multi-agency risk management (See chapter five). Within CDVP, there are sessions dedicated to the consideration of the impact of IPV upon children. Resistance to this sensitive material is to be expected and may be addressed more effectively where facilitators have an awareness of each group members' familial status and whether children were present during the offence. Encouraging perpetrators to personalise the material may promote change through increasing an awareness of the broader consequences to their behaviour. Additionally, utilising GLM principles to enable offenders to consider what they do well as parents may promote balance and reduce feelings of judgement or shame. Further research incorporating perpetrator,

partner and other family members' experience of behavioural change may provide a well-rounded view of treatment effectiveness.

A large number of the sample had histories of IPV representing a more entrenched pattern for many perpetrators. This is reflective of Chase et al. (1998) indicating that males may engage in a more relationship-specific pattern of offending. In the group room, this awareness may enable facilitators' to adopt a more patient and responsive approach, encouraging perpetrators to identify such patterns. Thames Valley Probation began to deliver a one-to-one version of CDVP for individuals requiring motivational work, increased support and an alternative option where group-work is not considered appropriate. This allows for flexibility of delivery which is supported other literature surrounding the varying risk levels and needs of an offender (Andrews & Bonta, 2010a). It is suggested from this research that a measure of IPV perpetration and victimisation are administered pre-treatment as this will highlight differences in the severity of perpetration amongst group members. This recommendation is based upon typology research (Johnson, 2008) and reflective of the 'need' principle of the RNR model.

Further research should aim to overcome some of the sampling and methodological difficulties within this study. This includes increasing the number of recidivists and non-recidivists within the sample to enhance generalisability, utilising a more culturally representative sample (Gondolf, 2011), including partner reports, increasing the length of follow-up and examining factors relating to loss of cases. With regards to drop-out variables, information relating to number of sessions attended, reason for drop-out will provide crucial information relating to the profile of participants and reduce confounding variables. A matched non-treatment control group would enhance the experimental quality of treatment effectiveness studies.

Gondolf (2001) reported that IPV perpetration is most likely to occur within six months of program commencement if they are receiving the program in the community. Research into levels of probation supervision, willingness to disclose and consistency of responses to disclosures of IPV incidents may provide insight into mediating factors of re-offending, risk assessment and reconviction relating to IPV.

Conclusions

Through utilising a sample of 259 male IPV perpetrators who were mandated to complete the CDVP in Thames Valley Probation Trust, this study has highlighted some interesting findings in relation to attitudinal and behavioural factors and insight into demographic factors among this population. This included positive indications of improvements on a number of psychological variables. Indeed, treatment effectiveness encompasses more than reported psychological change (Bowen et al., 2008; Laws & Ward, 2011). Whilst IPV reconviction was lower for individuals who had completed treatment, the differences disappeared when the construct of recidivism is expanded to include all other alleged IPV offending.

Higher risk of harm and increased breach among non-completers distinguished non-completers and completers demonstrating consistency with RNR principles (Andrews & Bonta, 2010a), levels of compliance and motivation (McMurran & Theodosi, 2007). Factors of interest relating to differences amongst treatment completer recidivists and non-recidivists included age, avoidant attachment, perceiving threat to the exclusivity within a relationship and marital status and requires further exploration within the theoretical context of impelling and inhibiting forces that may promote re-offending or cessation (Finkel et al., 2012).

Continued investigation is needed to answer the question of ‘what works’ in IPV treatment reflecting the RNR model (Akoensi et al., 2013). Whilst widely popular, the RNR model is not without its critics. Critique’s largely stem from the Good Lives Model (GLM,

Ward & Maruna, 2007) which places firm importance on an individual being empowered to achieve the life they desire (Laws & Ward, 2011). Laws & Ward (2011) have highlighted the RNR model's tendency to focus upon both criminogenic need (i.e. the negative) as opposed to an individual's protective factors at individual, social, environmental and cultural level that interact to promote rehabilitation and desistance (Ward & Maruna, 2007). The context of this should be placed in the underpinnings of how and why an intervention 'works' and for whom (Bowen, 2011; Lösel & Schmuker, 2005).

The current study demonstrates some positive outcomes for CDVP. A new IPV intervention has recently been accredited in the UK; Building Better Relationships (BBR) and is currently being delivered in some probation areas. This programme continues to utilise the nested ecological model however it reflects advancements in knowledge in the field of IPV through incorporating the GAM (Anderson & Bushman, 2002; see Chapter five). As such it will be important for the development of BBR to reflect upon aspects of CDVP that supported this sample to report psychological change and yield of a low-base rate of IPV reconviction. Further, this study can be used to enhance existing literature and programme delivery relating to IPV treatment as outlined above. In summary, this includes facilitators and treatment managers having greater awareness of attachment styles of individuals, promoting the consideration of the context and pattern of their IPV within the group room and its link to emotional experience and expression. Furthermore, this study highlights the need for joined up multi-agency working so as to address the co-occurrence of IPV and child witnesses to this and recognition of individual differences pertaining to drop out to enhance effective risk management.

Chapter Four

Critique of the State-Trait Anger Expression Inventory 2nd Version (STAXI-2: Spielberger, 1999)

Rationale for Chapter Four

The **State-Trait Anger Expression Inventory 2nd Version (STAXI-2: Spielberger, 1999)** is a self-report inventory that aims to explore assessed anger experience and expression among individuals and is widely used across forensic, clinical and general populations in assessment, treatment planning, research and evaluation of treatment. The STAXI-2 is used as part of the standardised battery of psychometric questionnaires at the pre- and post-treatment stages of the CDVP which was evaluated in Chapter three. The empirical study in Chapter three found self-reported change following treatment completion on the STAXI-2 pertaining to reported improvements in self-control and reduction in trait anger. Furthermore, non-completers reported higher levels of state and trait anger. However, there were also many ties reported within the scale of State anger, suggestive of no change on this subscale post-treatment, which indicates the need for further investigation and understanding.

Reviewing the STAXI-2 was considered relevant to this thesis as it is one of the few measures that has been validated across different age groups, ranging from adolescents to older-aged adults and has been used to investigate adult and adolescent IPV (Foshee et al., 2001; Murphy et al., 2007). As highlighted in Chapter one, the role of anger within both adolescent and adult IPV is complex and not fully understood. This is largely due to the conflicting theoretical perspectives within the paradigm of IPV. However, there is an evidence base to support associations between the difficulties in emotional regulation and the perpetration of IPV among adult populations (Babcock et al., 2000; Eckhardt et al., 2008; Norlander & Eckhardt, 2005) and adolescents (Moffitt et al., 2001). Lack of constraint and

negative emotionality factors accounted for a large amount of the variance when considering longitudinal risk factors amongst adolescents engaging in delinquent and violent behaviour (Moffitt et al., 2001). Furthermore, anger expression styles have been found to mediate the relationship between witnessing parental violence and AIPV perpetration although gender differences are observed (Wolf & Foshee, 2003). When considering IPV using Finkel's 'I^3 Theory' (Finkel, 2007; Finkel et al., 2009), one is drawn to the importance of understanding the role of anger experience and expression as a violence-impelling force, and anger control as violence-inhibiting force. For example, adult IPV offenders are found to report higher levels of anger and negative anger expression in comparison to non-violent men although this difference is modest (about one half of a standard deviation; Murphy et al., 2007).Whilst many IPV offenders may not have pathological levels of anger (Murphy et al., 2007), those who do experience problematic levels of anger perpetrate at higher frequencies resulting in higher levels of physical injury (Eckhardt, et al., 2008; Holtzworth-Munroe, Rehman & Herron, 2000; Murphy et al., 2007). With these issues considered, psychometric measures that examine an IPV offender's emotional regulation are often utilised and included as a component of IPV intervention evaluation (Babcock et al., 2004). Acceptable standards of validity of such measures are important to enable accurate and realistic insights into treatment development, treatment effectiveness and risk management.

This review will explore the utility of the STAXI-2 within the paradigm of IPV in assessing anger levels among perpetrators through examining its psychometric properties and research evidence.

Background

The construct of anger has long been examined by researchers although there has been confusion in its theoretical, conceptual and measurement properties (Deffenbacher et al., 1996. Spielberger (1999) has defined anger as "a psychobiological emotional state or

condition marked by subjective feelings that vary in intensity from mild irritation to annoyance to intense fury and rage" (p. 1). This definition highlights the multidimensional view of anger. Spielberger, Krasner and Soloman (1988) explained this as the 'AHA! Syndrome' comprising of the emotion of Anger, Hostility and the anger expression style (Aggression). More generally, anger is understood to be separate from hostility and aggression, with the former being related to a negative attitude towards others (Spielberger, 1999) and the latter encompassing destructive behaviours.

Anger as an emotional state is accompanied by a range of physical cues, including the activation of the autonomic nervous systems (Spielberger, 1999). The physiological reactions triggered by the experience of high levels of anger have been linked with a number of physical ailments including a risk of elevated blood pressure (Gentry, Chesney, Gary, Hall & Harburg, 1982).

The emotion of anger can be a precursor to verbal and physical aggression and is influenced by a range of interactions between personal, cognitive (such as hostile attributions) and environmental factors (Anderson & Bushman, 2002; Campbell, 2006; Deffenbacher et al., 1996; Kroner, Reddon & Serin, 1992).

Indeed there has been considerable controversy regarding the relationship between anger and violence among criminal offenders. Loza and Loza-Fanous questioned the usefulness of targeting anger in the treatment of rapists and violent offenders. In their studies (1999a; 1999b), no significant differences were found between violent and non-violent offenders or violent offender and rapists in their self-reported levels of anger and levels of risk identified by actuarial measures of recidivism. The psychometric results were observed at a single point in time which could be of limited clinical value (Mela et al., 2008) nor did it contain predictive data making any strong conclusions equivocal. Mills and Kroner (2003)

found no relationship between self-reported anger and post-release performance but again, no intervention was utilised within the study. It is logical to consider that the role of anger may not be applicable in all cases of aggressive behaviour however there is some evidence to support the association between reduced self-reported levels of anger and reduced further aggressive behaviours, particularly when engaged in an anger management programme (Mela et al., 2008). The debate detailing the link between anger and aggression appears stronger when applied within the area of IPV research (Bowen, 2011). Interestingly, Campbell (2006) discussed the disparity between findings that males are over-represented in their use of more severe forms of aggression yet sex differences in research examining males' and females' experience of anger yield no such difference and on occasion highlight women reporting more intense levels of anger. This may be considered reflective of prevalence studies in IPV among adults and adolescents whereby females report being more likely to be injured (Archer, 2000; Leen et al., 2013). However, it is an important element to consider given the similar perpetration rates found in AIPV/IPV.

There has been much research into exploring the *intentions* behind aggressive behaviour highlighting the dominant 'social cognitive' focus within the paradigm of aggression. Spielberger (2006) posits that gaining information about the variation, intensity and duration of an individual's emotions provides essential information and can ultimately contribute to more effective treatment. Therefore targeting an offender's subjective experience and expression of anger is an important area for interventions to address. Recently, Ferguson and Dyck (2012) have criticised this emphasis and consider it important for theories to explore the biological factors, genetic predispositions, personality factors and the interaction between that and the environment.

The STAXI-2 was developed by Charles Spielberger in 1999 for two purposes: "1) to assess components of anger for detailed evaluations of normal and abnormal personality, 2)

to provide means of measuring the contributions of various components of anger to the development of medical conditions, particularly hypertension, coronary heart disease, and cancer" (p. 1). The STAXI-2 and its predecessor, the STAXI (Spielberger, 1988), are arguably the most widely used measures of anger within clinical and research settings.

Overview of the tool

The STAXI-2 is a 57-item self-report inventory that is mainly used to assess various areas of anger among individuals. The items within the STAXI-2 consist of "six scales (*State Anger, Trait Anger, Anger Expression Out, Anger Expression In, Anger Control Out, Anger Control In*), five subscales and an Anger Expression Index that provides an overall measure of the expression and control of anger" (Spielberger, 1999, p. 1). The measure is presented in three parts. Part 1 explores an individual's present feelings; Part 2 explores an individual's general feelings and Part 3 explores how an individual reacts to these feelings. Individuals rate themselves on a 4-point Likert scale for each part (Part 1: 1 = not at all to 4 = very much, Parts 2 and 3: 1 = almost never to 4 = almost always). The items are totalled to produce a raw score (using the corresponding response value) which is converted into T scores and Percentile scores. The interpretative guidelines suggest scores that fall between the 25th and 75th percentiles may be considered to fall within the normal range. Anything above the 75th percentile signifies clinically significant difficulties with the experience and/or expression of anger to a degree that may interfere with essential functioning (Spielberger, 1999).

There are no reverse scored items within this inventory. The guidance time for STAXI-2 completion is 12-15 minutes (Spielberger, 1999, p. 4). A rating sheet form is available as well as an option to use a scannable rating form which enables large group screening and research.

The author provides detail on the development and validation of the inventory, administering, scoring and interpretive information and guidelines of normed populations. Research is included throughout the manual which serves to enhance the administrator's understanding of interpretation and implications. The reading level of the STAXI-2 items is pitched at an average 11-12 year olds' literacy ability.

State and Trait Anger

Spielberger (1999) explains that individual differences in the experience of anger can be conceptualised as having two main components: 'state' anger and 'trait' anger. These two scales are measured by the STAXI-2. State anger (S-Ang) refers to the psychobiological emotional state of anger that is characterised by subjective feelings in a specific situation. The S-Ang scale within the STAXI-2 includes three subscales: feeling angry (S-Ang/F), feeling like expressing anger verbally (S-Ang/V) and feeling like expressing anger physically (S-Ang/P). The STAXI-2 measures the intensity of the current feelings in relation to the above subscales. This version of the tool expanded upon the S-Ang scale contained within the original STAXI (Spielberger, 1988) and now contains 15 items as opposed to 10.

Trait Anger (T-Ang) refers to the level at which an individual tends to perceive situations as annoying or frustrating. The 10 T-Ang items in the STAXI-2 measures the frequency that general angry feelings are experienced across situations with higher levels of trait anger indicating that an individual has a disposition to view a wide range of situations as anger-provoking. The T-Ang scale includes two subscales: Angry temperament (T-Ang/T; which measures the readiness of an individual to express their anger) and angry reaction (T-Ang/R; measuring the tendency to react angrily). It has been found that individuals who report higher levels of trait anger experience more intense feelings of anger (state anger) (Deffenbacher et al., 1996).

Anger Expression and Control

The STAXI-2 also measures how often angry feelings are expressed and controlled. This is important when assessing an individual's level and use of aggression and can inform a treatment plan. The Anger Expression (AX) scale includes two subscales: anger expression-in (AX-I: holding in or suppressing anger) and anger expression-out (AX-O: expressing anger towards others people or objects). The Anger Control (AC) scale has two subscales: anger control-in (AC-I: controlling suppressed anger by calming themselves down) and anger control-out (AC-O: preventing the outward expression of anger). The AC-I subscale is not part of the original STAXI (Spielberger, 1988) and was developed for inclusion specifically for the STAXI-2. There are 32 Anger Expression and Control items within this inventory. Research has shown that high AX-I scores can be associated with physical ailments (Spielberger, 1999). Campbell (2006) discusses that few studies have found differences amongst males and females willingness or ability to express their anger. However, males are more likely to express this anger directly through verbal or physical aggression whilst studies have indicated that women are more likely to express their anger with an uninvolved person or cry. Novaco, Ramm and Black (2004) describe that the high prevalence of angry feelings along with anger dysregulation is a "significant component of psychological distress in the clinical profile among many types of offenders" (p. 130).

There is an overall Anger Expression Index (AX Index) that estimates the likelihood and frequency of expressing anger and comprising the AX-I, AX-O, AC-I and AC-O subscales. The formula with which this is computed is available in the manual. There is no Lie scale in this tool.

Psychometric Properties of the STAXI-2

Psychometric measures can provide an objective overview and evaluation of individuals' attitudes, behaviours and abilities. However, the quality of a measure's outcome is dependent upon the reliability and validity of the test itself. Further, its' psychometric properties should be available for its evaluation. The psychometric properties of the STAXI-2 have been validated through a range of studies (Spielberger, 1999). The STAXI-2 is an inventory and based upon measurement of responses to items that do not include a correct or incorrect answer (Coaley, 2009).

Reliability

A reliable psychometric test must enable consistent measurement of responses and be relatively free of measurement error (Kaplan & Saccuzzo, 2009). If this is not achieved, the measure cannot be considered a 'good test' (Kline, 1993). Spielberger and colleagues have conducted many studies to demonstrate the reliability and validity of the original STAXI (Spielberger, 1988). The published psychometric evidence for the STAXI-2 is more limited and whilst its' development was based upon empirical findings from the STAXI (Spielberger, 1988), there is scope for further research examining the robustness of the measure specifically given the revised structure and additional items within the STAXI-2.

Internal Reliability

Kline (1993) defines internal reliability as "the relationship of the items in a test" (p. 13) and therefore measures the consistency of a psychometric tool. Reliabilities for a good test should be .70 at a minimum (Kline, 1993).

Spielberger (1999) reported that the alpha coefficient measures of internal consistency for the STAXI-2 scales and subscales for the normal population and in-patient population were high across all scales and subscales (.84 or higher). The exception was on the T-Ang/R subscale for normal adult males and females which was .73 and .76 respectively (Spielberger, 1999). These results were based on a total sample size of 1664 individuals; 667 males and 997 females. Spielberger (1999) reported that these were satisfactory and not influenced by either gender or psychopathology. Spielberger presented further satisfactory alpha coefficients for the different age categories and the normal population. Alpha coefficients for the AX Index ranged from .75 to .82 showing satisfactory internal consistency. The highest co-efficient was found to be .95 (S-Ang scale; 16-19 year olds).

Kline (1993) asserted the importance to recognise that a very high internal consistency could suggest that all items are highly correlated may be indicative that the inventory is too specific in nature, with the items exploring the same question. As such, a very high alpha coefficient may limit test validity. That said, the level of the internal consistency found within the STAXI-2 suggests that the items are related in what they measure. Many researchers argue that high internal consistency between items is necessary (not sufficient) for good tests (Kline, 1993).

Test-Retest Reliability

Kaplan and Saccuzzo (2009) stated that test-retest reliability estimates are used "to evaluate the measurement of error associated with administering a test at two different times" (p. 109). Kline (1993) suggests that test retest reliability is an essential feature for any psychometric with .70 being the preferred standard of correlation to demonstrate this satisfactorily.

The test-retest reliability of the original STAXI was examined using different populations including undergraduate students (Jacobs, Latham & Brown, 1988) and incarcerated male offenders (Kroner & Reddon, 1992). The stability of the State scale was found to be higher than the Trait scale in a sample of 94 adult male offenders prisoners after one month; .88 and .64 respectively (Kroner & Reddon, 1992). Given that the State scale represents a more transient concept of fluctuating emotional experience in comparison to an individual's more stabilised traits, this finding was unexpected. Further, the researchers found that the test-retest stability was higher at the one month interval than one week (Trait scale coefficient, .57; State scale coefficient, .70).

Overall, research suggests satisfactory stability and it might be assumed that the test-retest reliability would be of a similar standard given that the STAXI-2 includes all of the items from the STAXI (plus additional items). However, there is no evidence within the manual that the STAXI-2 was re-administered to any of the standardised sample after a period of time. More research is required on a wider range of populations, particularly forensic samples, and over longer periods of time to provide a clearer picture of the test-retest reliability which incorporates the additional scales of the STAXI-2.

Validity

The validity of a test can be defined as the answer to the question: Does the test measure what it is supposed to measure? (Kaplan & Saccuzzo, 2009). They further describe the importance of obtaining evidence to show that the test has specific meaning which is observed in the findings of the test's relationship to other variables.

The STAXI-2 can be described as a standardised measure, because it can be administered by anyone and the results remain objective. A psychometric measure is said to

be standardised when comparisons can be made between individuals and populations (Coaley, 2009). Standardisation should be considered in relation to issues related to any limitations of the measure.

The validity of a test can be affected by the professionalism and the qualification of the administrators through misinterpretation. The STAXI-2 manual addresses this by stating that the test should be interpreted by a relevant professional who holds an "understanding of the concepts of anger, hostility and aggression" (Spielberger, 1999, p. 3).

Given that item responses are on a Likert Scale, the STAXI-2 questionnaire can be considered to have an ordinal level of measurement. This further enables more subtle effects to be explored using statistical analysis.

Face Validity

Kline (1993) refers to the face validity of a test as relating to the level at which it appears to be measuring what it claims to measure. It is the mere appearance of a test that demonstrates its face validity. Kaplan and Saccuzzo (2009) warn against the use of face validity as evidence of a valid test because it is not possible to objectively or directly measure it. If the items are seen to be related to what is being measured, a test user is able to understand and recognise the purpose of the tool. Positively, face validity may increase motivation to complete as individuals know what is being measured. In contrast, high face validity denotes transparency thus making it highly susceptible to faking or deliberately distorting their responses known as response bias.

An individual may attempt to present themselves in a favourable (socially desirable responding; faking good) or negative light (exaggerating the negative; faking bad). The effect

of this may become more apparent when a measure is solely reliant on self-report, as is the case with the STAXI-2 and particularly among offender populations (Foley, Hartman, Dunn, Smith & Goldberg, 2002).

Self-report measures exploring anger and aggression may be increasingly susceptible to response bias as a participant could be concerned about the negative consequences when disclosing aggressive incidents/ feelings. Individuals may attempt to deny or conceal their behaviour (James et al., 2005). Given the stigma and lack of open reporting found within the field of IPV (Saunders, 2008); face validity may pose more of a threat within this context. Although, it could also be argued that institutionalised populations may be willing to self-report their levels of anger and related cognitions and interpersonal difficulties as individuals may perceive that they have already experienced negative outcomes through sentencing.

The extent to which an individual is self-aware of feelings of anger and aggression will influence self-report measures. A problem arises when an individual may have come to view this as the norm (Bandura, 1986) thus arguably being unaware of any difficulty within the area of angry feelings. Spielberger (1999) suggests caution be taken with scores below the 25^{th} percentile in the STAXI-2 as this may be indicative of defensive responding.

Novaco et al. (2000) highlight the importance of inter-rating as opposed to gaining information through self-report alone. In relation to IPV, Archer (1999) demonstrated moderate correlations of .50 between participants self-report of physical incidents and their intimate partner's account. This method enhances congruent validity. The author is aware of the potential difficulties in clinical settings and implications for this method of administration. Similarly, it is important for other methods of assessment to be used in conjunction with self-report measures such as structured professional judgements, observations and reviewing background information.

Despite these difficulties, self-report within psychometric tests remains one of the most widely used largely due to the ease of administration, cost-effectiveness and time-effectiveness, however, many other measures include validity scales to take into account social, desirability responding (for example the Millon Clinical Multiaxial Inventory-III: MCMI-III, Millon, Millon, Davis & Grossman, 2006). In addition, questions aimed at an individual's real life experience are argued to yield a more accurate response in comparison to hypothetical situations (O'Connor, Archer & Wu, 2001).

While the STAXI-2 is regarded as having high face validity, it does not encompass a 'social desirability' scale and therefore the measure remains vulnerable to response bias despite its strong psychometric properties (Foley et al., 2002). In the case where the STAXI-2 is utilised within a forensic population, it is has been recommended that further validity scales of this nature are included within the STAXI-2 (Mela et al., 2008). This would enable the likelihood of a respondent's impression management to be scored, interpreted and give a more accurate view.

Concurrent Validity

A psychometric test is said to show concurrent validity if it can be evidenced to "correlate with another test of the same variable which was administered at the same time" (Kline, 1993, p. 17). Kline (1993) suggests .75 as a good correlation for concurrent validity but also highlighted that strong correlations could be indicative of two tests measuring the same thing, providing evidence that the test is not required. The STAXI-2 manual outlines research conducted using the original STAXI exploring the concurrent validity.

Concurrent validity of the original STAXI was evaluated (Spielberger, 1988) and moderately high correlations were found with the Buss-Durkee Hostility Inventory (BDHI,

Buss & Durkhee, 1957) and T-Ang scale in the STAXI. In addition, significant correlations were identified with Minnesota Multiphasic Personality Inventory Overt Hostility Scales (Schultz, 1954) and Hostility scales (Cook & Medley, 1954). Whilst this demonstrated evidence for the concurrent validity of the T-Ang scale as a measure of both anger and hostility in the original STAXI (Spielberger, 1999), it is of note that concurrent validity has not yet been established for the STAXI-2 and measures of hostility. The MMPI has since undergone significant normative (MMPI-2; Butcher, Dahlstrom, Graham, Tellegen & Kaemmer, 1989) and psychometric reconstruction (MMPI-2- Restructure Form; Ben-Porath & Tellegen, 2008/2011) with further research required to evaluate the concurrent validity of the STAXI-2 in this domain.

Moderate correlations were found with the Neuroticism domains in the Eysenck Personality Questionnaire (EPQ; Eysensck & Eysenck, 1975) and the T-Ang scale (Spielberger, 1988). In relation to the STAXI-2, small but significant positive correlations were found with the EPQ Psychoticism scale and State and Trait anger subscales in addition to low to moderate significant correlations between the EPQ Neuroticism scales and State (females = .27, males = .43) and Trait anger subscales (females = .49, males = .50) (Spielberger, 1999).

Culhane and Morera (2010) studied the validity and reliability of the STAXI-2 and Novaco Anger Scale (NAI; Novaco, 1994) among Hispanic and Non-Hispanic White student samples in North America. They found that the majority of scales had significant correlations in the expected direction when compared with the Multidimensional Anger Inventory (MAI; Siegel, 1986). Further evidence for the STAXI-2 includes findings of concurrent validity between the Swedish adapted version of the STAXI-2 and the NAS (Lindqvist, Daderman & Hellstrom, 2003). Both studies utilised a college sample and in order for sound conclusions to be made, further research is recommended particularly among offender populations.

The issue of response bias associated with the STAXI-2 among forensic populations (discussed above) was highlighted as potentially decreasing the concurrent validity of the measure (Foley et al., 2002).

Predictive Validity

Kline writes "A test may be said to have predictive validity if it will predict some criterion or other" (1993, p. 19), i.e. it is the extent to which a psychometric test may predict a future outcome. The predictive ability of the STAXI-2 within an offender population is a key element (i.e. its ability to predict potential recidivism).

In relation to the STAXI, there has been much research into the predictive validity of the measure in the medical field including chronic anger being associated with hypertension and high blood pressure (Spielberger, 1999). Deffenbacher et al. (1992) proposed that the original STAXI is an effective tool for screening individual's for anger treatment and also evaluating the treatment. This is supported by various studies (See Spielberger, 1999) although these use community samples. There has been less research among forensic populations.

Deffenbacher et al's (1996) studies among a student sample supported the utility of the T-Ang scale in predicting anger related incidents with correlations between AX and T-Ang being significantly higher than correlations between AX and anxiety and depression. When exploring the relationship between anger and offending, it appears that more research has explored the predictive validity of the STAXI as opposed to the STAXI-2.

Mela et al. (2008) conducted a 15-year outcome analysis study with 285 psychiatric offender patients in Canada who underwent an intensive anger management program. They were assessed pre- and post-treatment and used institutional data detailing anger-related

offending was obtained. They found that there was a significant decrease in the Anger Expression Index score and in the T-Ang scale post-treatment and that this was accompanied by a decrease in institutional offending. They concluded that the STAXI, along with offending data, is a useful assessment tool. On the other hand, Mills and Kroner (2003) found no significant relationship between anger and post-custody release success or failure (breach, further offence) following intervention.

It is of note that significant findings appear to be more consistent among a non-offending community based population. Foley et al. (2002) argue that more extensive research needs to be conducted within an offender population. Specific to IPV, Murphy and colleagues (2007) raised questions pertaining to the predictive ability of anger in the perpetration of IPV as the measure does not examine specific situational or attitudinal factors that may be relevant to partner abuse (see Chapter five). Methodological limitations within the above studies must be considered and further research is required before solid conclusions can be made in relation to offender populations and the STAXI-2.

Content Validity

In evaluating content validity, an attempt is made to determine whether a test has been constructed adequately and that the items included in the psychometric measure are relevant. Content validity is viewed as a more detailed form of face validity (Kline, 1993). It is also similar in that it is based upon logical evaluation as opposed to statistical (Kaplan & Saccuzzo, 2004). Importantly, clinical judgments are influenced by the construct validity of the psychometric tools which provide the information upon which the judgements are based.

Evaluating the content of the psychometric tool requires careful consideration of the appropriateness of each item ensuring that test items do not fail to capture elements of the

construct. A corroboration of expert judgements is viewed to be a more objective way of identifying the relevance of the items. Eckhardt, Norlander and Deffenbacher (2004) highlight that the factor analytic studies of the STAXI-2 (Spielberger, 1999) largely support the revised item and scale structure although there is some degree of item overlap across the feeling angry (S-Ang/F) and expressing anger verbally (S-Ang/V) factors in the S-Ang scale.

The outcomes of a psychometric tool may also be impacted by issues unrelated to the construct being measured such as reading ability, outside pressure, anxiety, stressful events and response bias. A factor that could influence completion of the STAXI-2 is the repetitiveness of the items, for example: "I feel like cursing out loud" and "I feel like swearing". This is linked to the transparency of the STAXI-2 and an individual's motivation to complete the task. Kline (1993) suggested that respondents may be inclined to avoid choosing the "a lot" option in relation to disclosing frequencies of behaviour. Content validation can be viewed as further evidence about the construct validity of an assessment instrument. Subsequently, this will be discussed further in the Construct Validity section.

Construct Validity

Kline (1993) considers that construct validity "embraces validity of every type" (p. 24). If a psychometric tool is 'construct valid', items will correlate with other tests measuring anger and will discriminate differences among groups. The STAXI-2 would therefore need to differentiate between individuals who experienced difficulties with experiencing, expressing or controlling anger from those who do not. Kline (1993) states that evaluating the construct validity of a test requires that the constructs being measured within the test must be clearly and operationally defined (Kline, 1993, p. 23).

The STAXI (and STAXI-2) was designed to assess anger in accordance with state–trait personality theory and to also further the understanding of the multiple components of

anger that are distinct from the constructs of hostility and aggression (See 'overview of tool' section). The theoretical distinction between these concepts is considered helpful in research.

Numerous structural factor analysis studies have tested the hypotheses in relation to the state-trait anger theory and provided consistent results for the construct validity of the STAXI (Fuqua et al., 1991; Spielberger, 1999). Deffenbacher et al. (1996) evaluated hypotheses emerging from the state-trait personality theory and showed support for the theory and the STAXI's measurement of the construct.

Extensive research has been conducted on the construct validity of the STAXI in the medical, physiological and educational fields. Within the field of psychology, studies have been carried out to examine the STAXI's construct validity in the assessment of anger in college populations (Deffenbacher et al., 1996), adolescent inpatients (Swaffer & Epps, 1999) and adult clinical inpatients in several countries including Brazil (Azevedo, Wang, Goulart, Lotufo, & Benseñor, 2010).

Spielberger (1999) proposes that the STAXI-2 measure has good discriminative ability yet readily acknowledges that this ability is reduced when attempting to discriminate between individuals at the lower end of the scale on the S-Ang and T-Ang/T scales. On the contrary, low scores on other scales such as AX-O and AX-I may provide useful information that is linked with potential denial or repression when managing anger.

Mela et al. (2008) highlighted that "if test scores do not translate into reduced anger-related offending, the utility of the tool may be brought into question" (p. 397). Foley et al.'s (2002) study found that only identified half of their 'engaged in anger treatment' sample as experiencing an anger problem to a level that would warrant treatment despite all subjects being engaged in treatment. They subsequently questioned the utility of the 75th percentile cut off point as the original STAXI did not effectively identify individuals with anger

problems. The researchers highlighted the need for further validity of the tools' use within an offender sample. Mela et al. (2008) went on to advocate that the utility of the 75th percentile cut off may lie within the expression scales of the STAXI as opposed to the experience scales, given that their sample did not score above the 75th percentile on S-Ang but did so on the Anger Expression (AX) subscales. With this in mind, it is important to review changes in AX and AC scales via the T-Ang scale, as suppression of anger can also have negative consequences both physically and psychologically (Spielberger, 1999).

Kroner and Reddon (1992) explored the reliability of the original STAXI with an incarcerated sample and their results questioned the value of the State-Trait distinction. However, further research has found Trait anger to be a well-known correlate of male-perpetrated physical and psychological aggression within an IPV context (Norlander & Eckhardt, 2005).

Within the domain of construct validity, Kline (1993) commented upon the discriminatory power of a psychometric measure; "its ability to produce a spread of scores" (p. 28) and thus have an ability to identify differences among individuals within and across populations. Within the field of anger and IPV offenders, Barbour, Eckhardt, Davison and Kassinove (1998) found that while IPV violent males may score higher on the T-Ang scale, they appeared to have difficulty in expressing their anger using emotional language and move to verbally aggressive behaviours. This, among other research, places support for STAXI's discriminatory power and also claims that domestically violent men experience higher levels of anger in a given situation, are predisposed to perceive an event as frustrating and also negatively express their angry emotions. However, this must be considered within the methodological limitations of such studies, namely small sample sizes which reduce the ability of such findings to be generalised to the wider IPV population. Further, Norlander and Eckhardt (2005) stated that the understanding of function of anger, hostility and IPV remains

elusive whilst Murphy et al. (2007) called for recognition of distinct anger profiles among IPV men due to the complexity of the role of anger and IPV perpetration.

The T-Ang and Anger Expression subscales within the STAXI-2 allow for further differentiation to be made when exploring different subtypes of IPV perpetrators (Barbour et al., 1998; Holtzworth-Munroe et al., 2000). Spielberger (1999) identifies characteristics of high trait anger and low anger control as linked with impulsive anger and high trait anger and a high level of anger control as being indicative of an individual potentially using anger in an intimidating manner. This is consistent with IPV typology research (Holtzworth-Munroe, et al., 2000; Norlander & Eckhardt, 2005) with evidence demonstrating that the most severely violent men have higher levels of anger (Holtzworth-Munroe et al., 2000) and are associated with higher levels of behavioural and interpersonal difficulties (Murphy et al., 2007). This is an area that would benefit from further examination both in relation to further validating the construct of the STAXI-2 among IPV subtypes and also the implications for responsivity issues within treatment.

A possible confounding factor with regards to differentiation is that the labelling of anger is a psychologically subjective process. The STAXI-2 asks individuals to rate whether they feel “annoyed” and also “furious”. These terms may be viewed by some as synonymous and makes an assumption that those completing the test can differentiate between varying intensities of anger. This wording could influence a way in which a person responds to the item.

Normative Samples

Norms refer to the observed performance of defined groups on a particular test and can give information about the performance relative to a standardised sample (Kaplan & Saccuzzo, 2009). The STAXI-2 manual contains normative data based on the responses of 1,920

individuals from two populations: normal adults (male and female) and for hospitalised male and female psychiatric patients. The mean age of the normal adults was aged 27 years, ranging from 16 to 63 years. No norm data has been collected for individuals under the age of 16 years. The age normed categories are: 16-19 years, 20-29 years and 30 years and older (Spielberger, 1999).

The substantial difference between the sample sizes (hospitalised psychiatric population = 276, community population = 1,644) means that caution should be taken when evaluating findings amongst psychiatric individuals particularly given the reported differences identified between them (Spielberger, 1999). This included higher levels of State and Trait anger, more frequent suppression and lower anger control amongst the psychiatric patients. Gender differences were also found with normal females having the lowest S-Ang scores and patient females having the highest. Specific gender differences were found that are consistent with the general expectation that males are more likely to express their anger and less likely to control their suppressed anger. It was found that overall, anger expression decreased with age and anger control increased. The findings support the need to apply age and gender appropriate norms when measuring and interpreting assessed anger, although this distinction should not be overstated (Campbell, 2006). Positively, the normed data has been separated by age and gender which enables a more robust interpretative ability of results in comparison to other psychometric measures of anger.

Studies into the utility and validity have tended to be focussed on college samples (Deffenbacher et al., 1996), inpatient hospitals (Mela et al., 2008), prisons (Kroner & Reddon, 1992) or mental health community settings (Foley et al., 2002). There have been studies that have used a community-based offender sample (Murphy et al., 2007) when examining treatment outcome and characteristics of IPV offenders. Given the increasing research relating to role of anger and its relationship to offending behaviour (Norlander &

Eckhardt, 2005), it is surprising that there are no standardised norms identified for a forensic population; prison- or community-based. The current psychiatric population norm is based upon a sample of inpatients undergoing treatment for psychiatric problems and addiction (Spielberger, 1999). It could be argued that these individuals may characteristically differ in comparison to violent and/or non-violent offenders. A standardised sample comprising violent offenders would be highly informative (Foley et al., 2002).

Use across settings

The STAXI-2 has been used across a range of settings to explore the relationship between anger and health. Hambleton (1994) has commented upon the issues surrounding the difficulties associated with the cross-cultural adaptation/ interpretation of psychological tests. These have included use of language, methodological errors and ensuring consistent connotations of an item. Such error is prone to occur when assessing the experience of emotions, given that they are subjective and less clearly defined (Anastasi, 1988). More specifically, the emotion of anger has been found to be culturally sensitive (Chon, 2002).

The STAXI-2 is primarily applicable to North American populations due to the standardised norms. The STAXI-2 continues to be adapted across a range of cultures and languages, enhancing its applicability to a wider population. This has led to the STAXI-2 being translated into Spanish (Miguel-Tobal, Casado, Cano-Vindel, & Spielberger, 2001), Swedish (Lindqvist et al., 2003) and Portugese (Azevedo et al., 2011). These adaptations represent western populations which could have implications when considering the attitudes and values that specific subcultures adhere to and populations outside of the western culture. More recently, research has examined the validity of the test across cultures outside of the Western world including Chinese (Maxwell, Sukhodolsky & Sit, 2009) and Iranian (Khodayarifard, Spielberger, Lavasani, & Zardkhaneh, 2013).

The positive results linked to the international use of the STAXI-2 demonstrates its adaptability across countries, however, it is a limitation that there have been no norms established as of yet within these countries. Caution should continue to be exercised when interpreting the results as the norms currently reflect a North American population. Assumptions should not be made about the construct of anger within different cultures. Studies that have examined the STAXI-2's applicability across settings are far fewer than for its predecessor.

With regards to specific ages, a Child and Adolescent version has been developed with its psychometric properties evaluated (del Barrio, Aluja, & Spielberger, 2004; Swaffer & Epps, 1999). This further demonstrates that the STAXI-2 possesses strong psychometric properties across a wide variety of normative groups (Culhane & Morera, 2010; Eckhardt, et al., 2004).

When investigating high school-aged dating violence, O'Keefe (1997) found that both males and females chose 'anger' as the most common reason for AIPV perpetration. Despite such findings, the evaluation of AIPV intervention programmes (e.g. Safe Dates Programme; Foshee et al., 2005) have largely focussed upon attitudinal change relating to the acceptability of dating violence to indicate positive change. However, attitudes form one part of a complex puzzle (see Chapter two) and indeed attitudinal change does not automatically translate to a reduction in aggressive behaviour as Shorey et al. (2012) point out. Utilising this age-appropriate version of the STAXI may provide more insight into the relationship between attitudinal change, reported anger experience and anger expression to further inform intervention and risk management of AIPV.

Research is increasing in exploring the role of anger among female perpetrators of IPV. Shorey, Brasfield, Febres and Stuart (2011) found that trait anger and impulsivity were

significantly associated to aggressive behaviour and that trait anger mediated the relationship between impulsivity and aggression among female perpetrators of AIPV (N = 80). This, and the above, indicates a requirement for continued research to enhance conclusive support for the use of the STAXI-2 with an offender population, more specifically its link to offending behaviour (Foley et al., 2002; Kroner & Reddon, 1992) and among female offenders.

Conclusions and Recommendations

This critique examined the psychometric properties of the STAXI-2 with respect to its reliability and validity. Research was discussed to explore the usefulness of the STAXI-2 (and original STAXI) in the assessment of anger as well as its use within research and practice.

Overall, the STAXI-2 has been found to be a good standard psychometric tool and its popularity and use among varying cultures cannot be disputed, particularly when applied to a normal population. It is recommended that further research is conducted exploring the test-retest reliability and concurrent validity specific to the STAXI-2 in light of advancing research findings. The validity of the tool has been subject to debate particularly in relation to its transparency. There should be further exploration into methods to minimise potential response bias associated with the STAXI-2 to further enhance its validity. This could include either simultaneously administering a robust ‘impression management’ measure or developing an embedded “Lie” scale within the STAXI-2 itself. Additionally, using various sources of information including social history, behavioural observations and inter-rater methods are important to draw robust conclusions and recommendations when the STAXI-2 is being utilised among an offender population.

There are many theories relating to anger, which present various challenges in measurement and construct. It could be argued that the STAXI-2 cannot take into account all aspects. There are of course discrepancies within the theoretical literature and evidently a psychometric tool will be constructed based upon the chosen theory and its range will be limited to that theory. Specifically, Eckhardt et al. (2004) raise the point that there is little known about measuring and understanding "clinical anger" as opposed to a personality dimension. Despite such limitations within the field of anger, the STAXI-2 is viewed to be based upon a solid conceptual model (Eckhardt et al., 2004) with clear definable boundaries. Professionals using this measure should do so in consideration of the strengths, limitations and debates in relation to the construct of anger.

The STAXI-2 has a range of norms, making it applicable to use within a range of cultures, ages and client groups and this critique has highlighted studies that have both supported, and questioned, the utility of the STAXI-2 among offender populations. Whilst there is generally consistent evidence of the higher levels of reported anger among perpetrators of IPV, the link to specific incidents and future offending remains elusive. Further research should focus on the application of the STAXI-2 pertaining to the predictive ability of this measure to behavioural outcomes such as further offending and the accuracy of the 'clinical cut-off' scores amongst offender populations. Given that the STAXI-2 continues to be widely used in the assessment, risk management and evaluation of treatment effectiveness among offenders, standardised norm data based on offender populations on which to interpret the STAXI-2 are crucial. In practice, those administering the measure to this population should ensure a clear explanation of the purpose of the measure and the potential benefits that can occur following open responding.

The English version of the STAXI and STAXI-2 are now becoming outdated as it has been in use since the early 1990's. Given the continuing development within the domain of

anger, it may be that more up to date tools are used in its place. Despite some criticisms to the STAXI-2, it remains a widely used psychometric measure and continues to play a role in much-needed research in exploring the multi-dimensional causes, experiences, expressions of anger.

Chapter Five

General Discussion

This thesis aimed to further understanding of intimate partner violence treatment effectiveness and wider practice issues across the lifespan. The focus was to pull together research from both adolescent and adult-aged populations to provide a well-rounded perspective of IPV throughout the life-course. This was done through three areas of study; a systematic literature review, a critique of a psychometric measure used in evaluating the Community Domestic Violence Programme (CDVP) within the UK; and an empirical piece of research exploring attitudinal and behavioural change relating to the completion or non-completion of CDVP.

The limitations within each chapter have been discussed and will naturally impact upon any conclusions and recommendations. Whilst it is important to consider the findings within the context of such limitations, the results this from research has practical implications for the field of AIPV and IPV with regards to prevention, treatment and risk management. Recommendations for future research discussed below in the format of a summary of findings relating to each chapter and subsequent discussion of the unifying themes from this thesis.

Summary of Findings

The first chapter in this thesis discussed the topic of IPV from a lifespan perspective (Theobald & Farrington, 2012), highlighting issues pertinent to IPV perpetration, victimisation and intervention among adolescents and adults. Definitions, prevalence, typology research and theoretical models of AIPV and IPV were discussed. Multifactorial frameworks including Dutton's Nested Ecological Model (1995; 2006), the GAM (Anderson

& Bushman, 2002) and I^3 Theory (Finkel, 2006) provided the backdrop for the complex nature of IPV perpetration and challenges faced when developing effective interventions.

Chapter Two presented a systematic literature review of studies that had examined the role of attitudes condoning violence among perpetrators and victims of AIPV (Gilbert et al., 2013). The aim of the review was to gain further understanding of the interplay between specific attitude types and the specific type of abusive behaviour used; physical, sexual or psychological/emotional (Holtzworth-Munroe, 2000; Johnson, 2006). The review began by highlighting the increasing prevalence of AIPV with adolescents not always viewing such behaviours within relationships as destructive (Schumacher & Smith Slep, 2004). The literature indicated a dearth of research in this area among high-school populations, particularly those in early adolescence (Wekerle & Wolfe, 1999).

After adopting a search criteria and quality assessment, seven studies were included. The review found some support that attitudes play a significant role in AIPV perpetration and victimisation, however, firm conclusions could not be made in relation to the authors' specific aims. This was due to the lack of investigation of the studies into psychological and sexual forms of abuse and the interchangeability of the definitions of 'attitudes'. Surprisingly, psychological abuse was excluded in some studies despite being identified as the most common type among this age group (Leen et al., 2013). There was also a lack of investigation into sexual AIPV with just one study including this type (Sears et al., 2007). This raised the issue that it is necessary for research to explicitly focus upon the different forms of partner abuse and violence. Practically, facilitators of intervention programmes should aim to explore the context behind sexual and psychological AIPV to enable distinction between partner-specific violence versus general interpersonal aggression. Chapter two discussed the need for further research into the development of healthy relationship attitudes from childhood to adolescence to adulthood.

An empirical paper examining the effectiveness of CDVP amongst a sample of adult male IPV offenders in the community was presented in Chapter Three. The aim was to contribute to the field of IPV given the continuing lack of conclusive evidence of existing treatment effectiveness research (Babcock et al., 2004; Hamilton et al., 2013). The study employed attitudinal and behavioural outcome measures, comparing treatment completers and non-completers and recidivists and non-recidivists. The findings of the study must be considered within its methodological limitations and sample sizes, however, some positive effects were found relating to psychometric variables and treatment completion. This included a reduction in; reported levels of jealousy regarding perceived threat to exclusivity to the relationship (Puente & Cohen, 2003; Stewart et al., 2005); trait anger and increased emotional control (Mela et al., 2008).

Overall recidivism rates were consistent with other studies (Rosenfeld, 1992; Stewart et al., 2005; Stover et al., 2009). Whilst IPV reconviction was lower for individuals who had completed treatment, the differences disappeared when the construct of recidivism is expanded to include all other alleged IPV offending. This raises issues around the disparity between official- and self-reports (Saunders, 2008) and questions the real-life transmission of cognitive change to behavioural change (Campbell, 2006). Chapter three discussed the implications for practitioners and it was raised that training and understanding of attachment styles, jealousy experience and expression and awareness of risk profiles is required when working with AIPV and IPV offenders.

Chapter four provided an in-depth critique of the psychometric properties of the STAXI-2 (Spielberger, 1999). This psychometric measurement was used as it forms one of the core measures of treatment change in the empirical study. The STAXI-2 is a standardised and widely utilised tool measuring an individual's assessed experience and expression of anger. It was considered to have strong psychometric properties including high internal

consistency and good level of reliability when applied to normal populations. It was recommended that efforts be made to establish a normalised sample for offender populations given its wide use in this field. The limitations of the tool related to its transparency and lack of a 'lie' scale as a means of counteracting this. This is particularly relevant with regards to offender populations (James, 2005). Inter-rating may be considered to increase the validity in this area, namely partner-reports for IPV offenders (Archer, 2000). A discussion regarding the construct and complexity of the interplay between anger and aggressive behaviour highlighted the need for further research specifically pertaining to IPV offenders (Norlander & Eckhardt, 2005). The application of anger profiles to IPV perpetrators similar to that found in typology literature was also highlighted (Johnson, 1995; 2006; Murphy et al., 2007).

Implications for AIPV/IPV prevention and intervention

For adult offenders who are mandated to complete IPV intervention within the National Probation Service, the predominant option continues to be to attend an accredited manualised group programme. Throughout this thesis, the heterogeneity of offenders with regards to age, gender, psychosocial variables and reasons for treatment drop-out has been observed. Adult non-completers presented with a higher risk profile (McMurran & Theodosi, 2007; Olver et al., 2011) and reported higher levels of jealousy (threat to exclusivity) and higher avoidant attachment (Allison et al., 2008). Attachment insecurity has been found to mediate the relationship between anger and IPV (Follingstad et al., 2002). Differences were also observed regarding extent to which adolescents held violence-supportive attitudes and experienced AIPV. Expectedly, 'high risk' community samples reported higher levels of this however such attitudes were not common among non-high risk community samples and adolescents held liberal gender role views.

Such differences pose challenges for treatment managers and facilitators, particularly when responding to the needs of individual group members. For example, having a comprehensive understanding of factors that may place an individual at high risk of drop-out at the pre-treatment stage may enable professionals to offer appropriate and additional support when an offender engages with such services. Whilst group-based treatment may be more cost-effective, the opportunity for individualised treatment plans should be provided to account for such heterogeneity. Individual interventions allow for more robust formulations and functional analyses of IPV behaviours (Bell & Naugle, 2008).

These findings must be also considered in the context IPV interventions particularly in exploring the dyadic nature in IPV occurs (Stith et al., 2012). Among adolescent-aged populations in the community, AIPV was largely bi-directional and was suggested to form more of a partner-specific pattern of behaviours for females, and relationship-specific for males (Chase et al., 1998). This is reflective of findings from Moffitt et al. whereby anti-social individuals tended to select mates with stronger anti-social profiles whereby couples engaged in bi-directional aggression (Moffitt et al., 2001). This is not to minimise male IPV perpetration but to highlight the importance in understanding the context of IPV, its triggers and the dynamics of an intimate partnership (Walker et al., 2013). Such findings have implications for intervention and prevention services (Archer, 2000; Shorey et al., 2012). Specifically, individuals who engage in a perpetrator programme may well have been victimised within their relationship and it is therefore important for practitioners, and researchers, to explore and understand these processes responsively. The majority of IPV intervention services within the UK are targeted at males as perpetrators, as is research. The findings from this thesis illustrate the importance of ensuring that female perpetration, and its' aetiology, is investigated using the same robust and reliable methods. Prevention and

intervention services for female perpetrators, adult and adolescent, should develop based upon robust sex differences, and similarity, research.

Despite attempts to examine younger-aged adolescents within this sample, the lack of studies meant that the majority of the conclusions and recommendations are based upon late-teens and adult populations. One study found that a sample of 11-12 years reported less healthy relationship attitudes and a 'double standard' whereby female-to-male violence was viewed as more acceptable than their older adolescent peers (Simon et al., 2010). This finding further highlight the importance of including younger-aged populations in research and prevention strategies. An accurate understanding of healthy relationships should be promoted and delivered through prevention strategies in schools. It is recommended that a focus on asking questions such as 'what is a healthy relationship?' as opposed to 'what is an unhealthy relationship?' reflective GLM principles (Laws & Wards, 2011), may serve to promote the benefits of having supportive intimate relationships rather than simply promoting 'avoidance of abusive and violent behaviour'.

This thesis highlighted that psychological change may not and should not be assumed to reflect changes in behaviour, i.e. self-reported behaviours or actual recidivism. This was evident within the systematic review (Leen et al., 2013; Shorey et al., 2012), the empirical study (Bowen et al., 2008) and the STAXI-2 critique (Mela et al., 2008; Mills & Kroner, 2003). Cognitive dissonance theory (Festinger, 1957) has proved useful in offering further understanding into the reduction of verbal aggression and jealous behaviour. Within a high school-aged sample, Schumacher and Smith Slep (2004) found that recognition and acceptance of cognitive dissonance between attitudes supportive of IPV and aggressive behaviours reduced from Time 1 to Time 2. This approach has been utilised in motivational interviewing (Miller & Rollnick, 2002) and treatment programmes within the UK but within its infancy of application to the IPV field. It is recommended that applying such theory to

prevention and intervention services may increase the link between attitudinal and behavioural change. In this context, group facilitators and practitioners should ensure that questions are asked relating to an individual's cognitions and emotions during skills-practices (role plays) as this will provide opportunities to recognise, reflect upon, any dissonance.

Evaluation measures presented as group means cannot inform change at an individual level (Bowen et al., 2008), and whilst helpful to provide overall trends in treatment effectiveness, it is recommended that individual change is evaluated, as observed in the methodology of Bowen and colleagues' (2008) study to provide a practically meaningful understanding as to the underlying mechanism of change.

Implications for policy and risk management

Across both populations, there remains continued difficulty in adopting operational definitions with policymakers and researchers often using different terms (CDC, 2008; Glass et al., 2003; see Chapter two). This has implications for the consistent and comparability of research within this field, namely in investigating, understanding and policy-making related to psychological IPV, sexual IPV and controlling and coercive behaviours. Positively, the cross-government definition has been extended to include 16-17 year olds (Home Office, 2013) demonstrating recognition of the seriousness of IPV that occurs amongst those under the age of 18 years. This extension does lead to questions pertaining to the age-ranges associated with the definition of AIPV and IPV. Indeed, investigating violence perpetrated against a partner across the lifespan is imperative within this field and similarly, distinctions between martial violence versus dating violence should also be made across age groups to further inform policy, prevention and intervention across different intimate partnership types (Shorey et al., 2008). This could also be extended to younger-aged adolescents whereby

dating relationships take on a different form and meaning to the individual (Simon et al., 2010).

Very recently, a change to UK government legislation now renders controlling and coercive behaviour within an intimate or familial relationship a criminal offence in the UK (The Serious Crime Act 2015). This aims to capture continuous non-violent abusive behaviours within a relationship not adequately addressed through harassment and stalking laws, particularly where a relationship is on-going. These changes should provide a clearer foundation for robust research to examine specific types of IPV across ages although time will be required to evaluate how this legislation change impacts research, policy, society and services for victims and perpetrators.

Among the adolescent samples within studies, the author was unable to ascertain whether the AIPV perpetrated by 'non-high-risk' samples was less, more, or similarly severe when compared to those in a 'high-risk' sample nor what the consequences to such behaviours were. Interestingly, the results of the empirical study indicated that whilst more treatment completers disclosed alleged IPV re-offending, more non-completers received IPV reconvictions. Whilst not statistically significant, it does raise important questions pertaining to the role of the OM, the opportunity for disclosure and consistency of responses to an alleged IPV re-offence by practitioners. Policymakers should consider the consistency of legal and social consequences in response to AIPV / IPV perpetration, ensuring risk assessment and appropriate management is at the fore (Howard, 2006).

Chapter three highlighted a large number of children were present during the perpetration of IPV offences. Witnessing violence in the home has been identified as a longitudinal risk factor for subsequent IPV perpetration (Lussier et al., 2009; Narayan et al., 2014; Roberts et al., 2010). Indeed, the intergenerational cycle is not as prevalent amongst community samples (Stith et al., 2000). Nevertheless, the use of AIPV during adolescence

has been linked to IPV perpetration as an adult (O'Leary, 1986) although this requires further investigation (Shorey et al., 2008). There is considerable overlap between the two age groups examined in this thesis regarding perpetration, victimisation and familial contexts of IPV which is supported by existing research (Dixon & Graham-Kevan, 2011; Ehrensaft et al., 2004; Moffitt et al., 2001; Stith et al., 2012; Theobald & Farrington, 2012). Therefore children and adolescents must be a consideration when addressing prevention, intervention and risk management of IPV.

Within the context of child maltreatment that may arise in witnessing IPV (Dixon & Graham-Kevan, 2011); the co-operation of multi-agencies such as child protection, schools, health and domestic violence agencies remains imperative. This includes the effective sharing of information regarding child safety following on from any disclosures from perpetrators or victims of further abuse respective of confidentiality and child protection protocol. This will allow for more appropriate and tailored support for family members. Agencies should be directly involved with the family and children where appropriate in relation to IPV issues. Practical examples of this may include; trauma-focussed child counselling, education and prevention efforts to reduce the likelihood of intergenerational transmission of anti-social and abusive behaviours or intervention for adolescents who may be engaging in AIPV perpetration in their own relationships. This will help to try and reduce the impact of IPV on the next generation (Ehrensaft et al., 2004; Moffitt et al., 2001).

Future Directions

This thesis has demonstrated support for a gender-inclusive multi-faceted approach regarding IPV perpetration by both adolescent and adult populations (Dutton & Corvo, 2006). This is evident in the unifying theme across the chapters that a single variable cannot fully

explain IPV perpetration (see Chapter two) nor does reported change on a psychological variable determine or distinguish recidivists from non-recidivists (see Chapter three). Indeed, the plethora of risk factors associated with IPV perpetration across the lifespan creates challenges of how to best understand, address and manage such behaviours (Finkel, 2007). Whilst positive outcomes were observed relating to psychological change in the empirical study, further understanding is required as to what inhibits and/or impels specific types of IPV amongst adolescents and adults. When considering the underlying mechanism of attitudes relating to IPV perpetration, the research has not succeeded in examining the depth of complexities. It is recommended, in addition to the consideration of cognitive dissonance theory (Festinger, 1957) mentioned above, that Finkel's I^3 Theory (Finkel, 2007; Finkel et al., 2012) is utilised in further research into etiology and treatment evaluation of IPV across all age ranges. Its focus upon the balance of violence-impelling factors, such as self-regulatory failure, versus violence-inhibiting factors, such as fear of consequences, at the point of a conflict will enable a more in-depth contextual understanding of dyadic processes (Stith et al., 2012).

Anger and anger expression literature suggest that dispositional anger should not be considered in isolation but as a complex interplay of factors including instigation, context and inhibition (Campbell, 2006; Finkel et al., 2012). The 'perfect storm' analysis used within Finkel's framework (i.e. strong instigation, strong impellance and weak inhibition) found that even those who are high on dispositional aggressiveness did not engage in physical IPV (Finkel et al., 2012). Thus suggesting that changing one of these elements may be effective in reducing the likelihood of IPV perpetration and victimisation and should be considered in IPV intervention. This theory could provide further understanding into the finding that the majority of in the empirical study within this thesis sample in this thesis did not meet the 75th

percentile cut off outlined in the STAXI-2 manual (Spielberger, 1999), consistent with Mela et al. (2008) and Murphy et al. (2007).

Whilst non-completers reported higher levels of trait anger (Barbour et al., 1998), the STAXI-2 variables did not distinguish recidivists from non-recidivists in the current study (Mills & Kroner, 2003). Further research regarding the construct validity of the STAXI-2 in distinguishing between those who have problems with anger and aggression among an IPV sample is required but it could be considered that even though some IPV offenders may not report issues with anger and aggression to the point of clinical significance, weak inhibition (such as low self-regulation or beliefs that IPV will not be punished) may indeed reduce an individual's ability to cope with even slight angry feelings and/ or aggressive impulses (Finkel et al., 2012). It is therefore imperative for practitioners and researchers to tap into the context in which such IPV incidents occurred and examine the strength of inhibition within IPV intervention whether individual-, group- or couples-based (Bell & Naugle, 2008). Furthermore, understanding and strengthening an individual's ability to employ dispositional, situational and relational violence-inhibiting skills within their intimate relationships is reflective of more recent desistance literature (Walker et al., 2013) and can utilise GLM principles (Ward & Maruna, 2007). This includes the role of motivation (Bowen & Gilchrist, 2006) which is not currently assessed before or after treatment in the CDVP using psychometric measures. Such information is crucial in identifying mediating factors relating to change in attitudes and functioning (Connors et al., 2012).

Evaluation of I^3 theory is in its infancy and at present has not yet been utilised amongst adolescent-aged samples. However, utilising more recent frameworks, based upon up-to-date research will serve not only to reflect RNR principles (Andrews et al., 1990) but afford a move away from the more ideological approaches to IPV aetiology and intervention (Bowen, 2011; Dixon & Graham-Kevan, 2012). It is important to note that advances are

being made regarding applying theoretical models to adult treatment interventions in the UK, in the form of the Building Better Relationships programme (BBR)., which incorporates the GAM and the Good Lives Framework (Ward & Maruna, 2007) in addition to the nested ecological model (Dutton, 1995; 2006) as in CDVP. Such advancements are not yet reflective within AIPV theories (Shorey et al., 2008) which is a direct requirement to ensure that development of interventions are evidence based (Andrews et al., 1990).

This thesis has examined IPV across two stages of the life-course; adolescence and adulthood. Further research should be dedicated to examining IPV amongst the older population utilising developmental theories and contextual frameworks with an aim to understand the applicability of them to the older population. Given that increasing life expectancy, it is important to gain understanding of the prevalence, stability, and context of this phenomenon to ensure that victims and perpetrators of all ages are appropriate supported.

Emotion and cognition are of course inextricably linked in complex patterns (Campbell, 2006) and what it clear from the work within this thesis is that multifaceted theories must be adopted to gain understanding of the differences, and similarities, pertaining to both gender and specific dyadic processes between couples. The critique of the STAXI-2 also identified that this measure may not adequately capture such complexities of anger within the context of AIPV/IPV. Furthermore, the methodological differences and simplification of the construct of attitudes relating to AIPV yielded inconclusive results regarding the differences in cognition regarding different types of the intimate partner violence studies examined in Chapter two. Moreover, evaluative measures do not adequately explore the context in which such beliefs and attitudes occur (Follingstad et al., 2002). Building a deeper understanding of such mechanisms will aid the development of effective prevention campaigns and treatment programmes for both adolescents and adults. In order to achieve this, qualitative methods could be utilised in examining situational triggers of

specific attitudes and IPV behaviours (Johnson, 2006). It is therefore recommended that research continues to ensure valid, reliable measures are used when evaluating interventions. Practically, facilitators working collaboratively with an individual at the pre-treatment stage to complete pre-treatment measures may make the process more personal and meaningful; empowering the individual to consider their strengths and needs.

The empirical study within this thesis utilised behavioural outcomes, strengthening its methodology by not solely focussing upon attitudinal measures as many treatment evaluation studies do. Whilst using recidivism measures is not without its critics and challenges (Saunders, 2008), the findings within this thesis have provided a useful exploration into the differences between treatment completers and non-completers, and recidivists versus non-recidivists. Examining recidivism, and desistance, should not be limited to physical IPV. Psychological IPV is the most common form of dating violence (Leen et al., 2013) and is included in the cross-government definition of IPV. However, there are few or no prevention or intervention programmes for this phenomenon specifically and should be addressed (Shorey et al., 2012). Further research is required to gain more understanding of the use of psychological aggression across the lifespan.

Within the context of IPV perpetration, the impact of the rapid developments in technology and social media are yet to be fully understood. Abusive behaviours such as publicly shaming an individual on a social networking site or a perpetrator sharing a sexually explicit video without a partner's consent need to be understood and addressed. It is imperative that practitioners, prevention services and policy-makers understanding of IPV across all age ranges reflect culturally relevant and present day issues. Moreover, services that aim to prevent IPV and promote the safety of victims must continually be updated to ensure it is meaningful to the individual it is targeted at and whom it is aiming to protect at all levels.

It is the aim of the researcher to conduct further analyses on the data within the current study to gain increased insight into IPV treatment effectiveness regarding individual change (Bowen et al., 2008), the strength of association between attitudinal and behavioural change and its' predictive ability amongst this specific sample of male offenders. Observational methods and inter-rating are suggested as ways of increasing accuracy and objectivity in assessing effectiveness. Further, research whereby an individual is monitored from the start of the programme and throughout will provide further insight into treatment effectiveness.

Conclusions

The progress in the development of IPV prevention and intervention services, policy and research continue to be hampered by the longstanding debate between gender exclusive versus gender inclusive approaches (Dobash & Dobash, 2004; Dutton & Corvo, 2006; Gondolf, 2001; Straus, 2008). Whilst research has flourished in an effort to understand the phenomenon of both AIPV and IPV, it has remained difficult to provide a broad framework that effectively incorporates the interplay of the dozens of evidenced risk factors (Finkel et al., 2012). Some scholars state this may lead to more problems in theory than in practice (Akoensi et al., 2013) yet despite this the field has struggled to respond to research findings with the effectiveness of treatment remaining inconclusive and modest at best (Babcock et al., 2004; Bowen, 2011; Shorey et al., 2012).

This thesis demonstrates the importance of improving awareness and understanding of all forms of intimate partner violence across the lifespan given the difficulties that are seemingly inherent in developing effective interventions within the field (Akoensi et al., 2013; Bell & Naugle, 2008; Leen et al., 2013; Shorey et al., 2012). Whilst research may have

identified what elements guide effective intervention (Andrews & Bonta 2006), adopting a holistic approach enables further understanding of the underlying mechanisms of how an intervention 'works' (Bowen & Gilchrist, 2004). BBR is delivered in some probation trusts and will eventually phase out CDVP and IDAP although CDVP continues to be delivered within the UK prison service (as HRP). The findings of this thesis can inform IPV treatment effectiveness more widely within the UK particularly as some of theoretical foundations remain the same for CDVP and BBR as highlighted.

Given the wide array of behaviours associated with IPV, it is imperative that academics researching this field respond to recent developments in the legislation and the definition of IPV within the UK, placing a focus upon investigating the perpetration and victimisation of psychological, sexual and controlling behaviours amongst adolescents and adult populations. Whilst the relevance of context and dyadic factors was highlighted within this thesis, research methodologies employing longitudinal sibling samples is encouraged to reflect holistic insight.

Whilst there are evidenced differences between marital violence and dating violence (Shorey et al., 2008), findings also indicate many of the same pre-disposing, mediating and moderating risk factors are relevant for both age-groups (Shorey et al., 2008). The development of a comprehensive framework, and application of current evidence-based theory, should continue guide academics and practitioners within the UK regarding causality and the developmental trajectory of IPV (Lussier et al., 2009). This will enhance the development of effective interventions at primary, secondary and tertiary levels of adolescent and adult-age populations (McGuire, 2013). Results and recommendations from research using UK based IPV samples can be more accurately applied to UK criminal justice policy and support services rather than relying on its' North American counterparts. Such research is vital in aiding public policy in how to work most effectively with offenders.

References

Akoensi, T. D., Koehler, J. A., Lösel, F., & Humphreys, D. K. (2013). Domestic violence perpetrator programs in Europe, Part II: A systematic review of the state of evidence. *International Journal of Offender Therapy and Comparative Criminology, 57,* 1206-1225. doi:10.1177/0306624X12468110

Ali, B., Swahn, M., & Hamburger, M. (2011). Attitudes affecting physical dating violence perpetration and victimization: Findings from adolescents in a high-risk urban community. *Violence and Victims, 26,* 669-683. doi:10.1891/0886-6708.26.5.669

Allison, C. J., Bartholomew, K., Mayseless, O., & Dutton, D. G. (2008). Love as a battlefield: Attachment and relationship dynamics in couples identified for male partner violence. *Journal of Family Issue, 29,* 125-150. doi:10.1177/0192513X07306980

Anastasi, A. (1988). *Psychological testing (6th ed.).* New York: Macmillan.

Anderson, C. A., & Bushman, B. J. (2002). Human aggression. *Annual Review of Psychology, 53,* 27-51. doi:0084-6570/02/0201-0027$14.00

Andrews, D. A., & Bonta, J. (2010a). *The psychology of criminal conduct (5th ed.).* New Providence, NJ: LexisNexis Group.

Andrews, D. A., & Bonta, J. (2010b). Rehabilitating criminal justice policy and practice. *Psychology, Public Policy and Law, 16,* 39-55. doi:10.1037/a0018362

Andrews, D. A., Bonta, J., & Hoge, R. (1990). Classification for effective rehabilitation: Rediscovering psychology. *Criminal Justice and Behaviour, 17,* 19-52. doi:10.1177/0093854890017001004

Andrews, D. A., Bonta, J., & Wormith, J. S. (2011). The Risk-Need-Responsivity (RNR) model: does adding the Good Lives Model contribute to effective crime prevention? *Criminal Justice and Behaviour, 38,* 735-755. doi:10.1177/0093854811406356

Archer, J. (1999). An assessment of the reliability of the Conflict Tactics Scales: A meta-analytic review. *Journal of Interpersonal Violence, 14,* 1263-1289. doi:10.1177/088626099014012003

Archer, J. (2000). Sex differences in aggression between heterosexual partners: A meta-analytic review. *Psychological Bulletin, 126,* 651-680.doi:10.1037/MJ033-2909.126.5.651

Archer, J. (2006). Cross-cultural differences in physical aggression between partners: A social-structural analysis. *Personality and Social Psychology Review, 10,* 113-133.doi:10.1207/s15327957pspr1002_3

Azevedo, F. B., Wang, Y., Goulart, Y. C., Lotufo, P. A., Benseñor, I. M. (2011). Application of the Spielberger's State-Trait Anger Expression Inventory in clinical patients. *Arquivos de Neuro-Psiquiatria, 68,* 231-234. doi:10.1590/S0004-282X2010000200015

Babcock, J. C., Green, C. E., & Robie, C. (2004). Does batterers' treatment work? A meta-analytic review of domestic violence treatment. *Clinical Psychology Review, 23.* 1023-1053. doi:10.1016/j.cpr.2002.07.001

Babcock, J. C., Jacobson, N. S., Gottman, J. M., & Yerington, T. P. (2000). Attachment, emotional regulation, and the function of marital violence: Differences between secure, preoccupied, and dismissing violent and nonviolent husbands. *Journal of Family Violence, 15,* 391-409. doi:10.1023/A:1007558330501

Baker, N. L., Buick, J. D., Kim, S. R., Moniz, S. & Nava, K. L. (2013). Lessons from examining same-sex intimate partner violence. *Sex Roles, 69,* 182-192. doi:10.1007/s11199-012-0218-3

Baker, C. R., & Stith, S. M. (2008). Factors predicting dating violence perpetration among male and female college students. *Journal of Aggression, Maltreatment and Trauma, 17,* 227-244. doi:10.1080/10926770802344836

Bandura, A. (1986). *The social formations of thought and action: A social cognitive theory.* Englewood Cliffs, NJ: Prentice Hall.

Barbour, C., Eckhardt, C., Davison, J., & Kassinove, H. (1998). The experience and expression of anger in maritally violent and maritally discordant -nonviolent men. *Behavior Therapy, 29,* 173-191. doi:10.1016/S0005-7894(98)80001-4

Barter, C., McCarry, M., Berridge, D., & Evans, K. (2009). *Partner exploitation and violence in teenage intimate relationships: A research report from the University of Bristol and NSPCC.* Available from https://www.nspcc.org.uk/Inform/research/findings/partner_exploitation_and_violence_wda68092.html.

Baumeister, R. F., & Vohs, K. D. (2004). Handbook of self-regulation: Research, theory, and applications. New York: Guilford Press.

Bell, K. M., & Naugle, A. E. (2008). Intimate partner violence theoretical considerations: Moving towards a contextual framework. *Clinical Psychology Review, 28,* 1096-1107. doi:10.1016/j.cpr.2008.03.003

Ben-Porath, Y. S., & Tellegen, A. (2008/2011). *MMPI-2-RF (Minnesota Multiphasic Personality Inventory-2 Restructured Form): Manual for administration, scoring, and interpretation.* Minneapolis, MN: University of Minnesota Press.

Bersani, C. A., Chen, H. T., Pendleton, B. F., & Denton, R. (1992). Personality traits of convicted male batterers. *Journal of Family Violence, 7,* 123-134. doi:10.1007/BF00978701

Bethke, T. M., & DeJoy, D. M. (1993). An experimental study of factors influencing the acceptability of dating violence. *Journal of Interpersonal Violence, 8,* 36-51. doi: 10.1177/088626093008001003

Blumstein, A., & Larson, R. C. (1971). Problems in modelling and measuring recidivism. *Journal of Research in Crime and Delinquency, 8,* 124-13. doi: 10.1177/002242787100800202

Bowen, E. (2010). Therapeutic environment and outcomes in a U.K. domestic violence perpetrator program. *Small Group Research, 41,* 198-220. doi:10.1177/1046496409357081

Bowen, E. (2011). *The Rehabilitation of Partner-Violent Men.* Chichester: John Wiley and Sons Ltd.

Bowen, E., & Gilchrist, E. (2004). Comprehensive evaluation: A holistic approach to evaluating domestic violence offender programmes. *International Journal of Offender Therapy and Comparative Criminology, 48,* 215-234. doi:0.1177/0306624X03259471

Bowen, E., & Gilchrist, E. (2006). Predicting dropout of court-mandated treatment in a British sample of domestic violence offenders. *Legal and Criminological Psychology, 12,* 573-587. doi:10.1080/10683160500337659

Bowen, E., Gilchrist, E. A., & Beech, A. R. (2005). An examination of the impact of community-based rehabilitation on the offending behaviour of male domestic violence offenders and the characteristics associated with recidivism. *Legal and Criminological Psychology, 10,* 189-209. doi:10.1348/135532505X36778

Bowen, E., Gilchrist, E., & Beech, A. R. (2008). Change in treatment has no relationship with subsequent re-offending in U.K. Domestic violence sample: A preliminary study. *International Journal of Offender Therapy and Comparative Criminology, 52,* 598-614. doi:10.1177/0306624X08319419

Brown, M. B., & Forsythe, A. B. (1974). Robust tests for the equality of variances. *Journal of the American Statistical Association, 69,* 364-367. doi:10.1080/01621459.1974.10482955

Bullock, K., Sarre, S., Tarling, R., & Wilkinson, M. (2010). *The delivery of domestic abuse programmes: An implementation study of the delivery of domestic abuse programmes in probation areas and Her Majesty's Prison Service.* Ministry of Justice Research Series 15/10. Available from https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/205168/delivery-domestic-abuse-programmes.pdf.

Buss, A. H., & Durkee, A. (1957). An inventory for assessing different kinds of hostility. *Journal of Consulting Psychology, 21,* 343-349. doi:10.1037/h0046900

Butcher, J. N., Dahlstrom, W. G., Graham, J. R., Tellegen, A. M., & Kraemmer, B (1989). *The Minnesota Multiphasic Personality Inventory-2 (MMPI-2) Manual for Administration and Scoring.* Minneapolis, MN: University of Minneapolis Press.

CAADA (2012). *Insights Factsheet: Teenage victims of domestic abuse.* Available from http://www.ccrm.org.uk/images/docs/4.2acaada-insights-factsheet-teenage-domestic-abuse-final.pdf.

Callahan, M. R., Tolman, R. M., & Saunders, D. G. (2003). Adolescent dating violence victimization and psychological well-being. *Journal of Adolescent Research, 18(6).* 664-681. doi: 10.1177/0743558403254784

Carney, M. & Dutton, D. (2007). Women who perpetrate intimate partner violence: A review of the literature with recommendations for treatment. *Aggression and Violent Behavior, 12, 108*–115. doi:10.1016/j.avb.2006.05.002

CDVP Manual. (2008). *Community Domestic Violence (CDVP) Manual.* National Offender Management Service (NOMS).

Centers for Disease Control and Prevention (CDC; 2008). *Understanding Teen Violence Factsheet.* Available from http://www.cdc.gov/ncipc/pub-res/datingabusefactsheet.pdf.

Chase, K. A., Treboux, D., O'Leary, D. K., & Strassberg, Z. (1998). Specificity of dating aggression and its justification among high-risk adolescents. *Journal of Abnormal Child Psychology, 26,* 467-473. doi:10.1023/A:1022651818834

Chon, K. K. (2002). Cultural aspects of anger. *XXVII International Congress of Psychology, 2,* 323-346.

Criminal Justice Act. (2003). *Schedule 15. Specified offences for purposes of Chapter 5 of Part 12.* Available from http://www.legislation.gov.uk/ukpga/2003/44/schedule/15

Coaley, K. (2009). *An Introduction to Psychological Assessment and Psychometrics.* London: Sage Publications.

Cohen, J. (1992). A power primer. *Psychological Bulletin, 112,* 155-159. doi:10.1037/0033-2909.112.1.155

Connors, A. D., Mills, J. F., & Gray, A. L. (2012). Evaluation of intimate partner violence intervention with incarcerated offenders. *Journal of Interpersonal Violence, 27,* 1176-1196. doi:10.1177/0886260511424499

Cook, W. W., & Medley, D. M. (1954). Proposed hostility and pharisaic-virtue scales for the MMPI. *Journal of Applied Psychology, 38,* 414-418. doi: 10.1037/h0060667

Culhane, S. E., & Morera, O. F. (2010). Reliability and validity of the Novaco Anger Scale and Provocation Inventory and State-Trait Anger Expression Inventory-2 in Hispanic and Non-Hispanic White student samples. *Hispanic Journal of Behavioral Sciences, 32,* 586-606. doi: 10.1177/0739986310381458

Davis, R. C., & Taylor, B. G. (1999). Does batterer treatment reduce violence? A synthesis of the literature. *Women and Criminal Justice, 10,* 69–93. doi:10.1300/J012v10n02_05

Deblinger, E., Weiss, A., Hoch-Espada, A., & Steer, R. H. (2000). *The relationship attitudes survey for adolescents: A measure of adolescent attitudes regarding sexuality and coercion.* Unpublished manuscript, UMDNJ-SOM, Stratford, NJ.

Deffenbacher, J. L. (1992). Trait anger: Theory, findings, and implications. In C. D. Spielberger & J. N. Butcher (Eds.), *Advances in Personality Assessment* (pp. 177-201). Hillsdale, NJ: Erlbaum.

Deffenbacher, J. L., Oetting, E. R., Thwaites, G. A., Lynch, R. S., Baker, R. S., Thacker, S., & Eiswerth-Cox, D. (1996). State -Trait Anger Theory and the utility of the Trait Anger Scale. *Journal of Counseling Psychology, 43,* 131-148. doi:10.1037/0022-0167.43.2.131

del Barrio, V., Aluja, A., & Spielberger, C. D (2004). Anger assessment with the STAXI-CA: psychometric properties of a new instrument for children and adolescents. *Personality and Individual Differences, 37,* 227-244. doi:10.1016/j.paid.2003.08.014

DeWall, C. N., Anderson, C. A., & Bushman, B. J. (2011). The General Aggression Model: Theoretical extensions to violence. *Psychology of Violence, 3,* 245-258. doi:10.1037/a0023842

Devine, J. A., Wright, J. D., & Joyner, L. M. (1994). Issues in implementing a randomized experiment in a field setting. *New Directions for Program Evaluation, 63,* 27-40. doi: 10.1002/ev.1681

Dixon, L., Archer, J., & Graham-Kevan, N. (2012). Perpetrator programmes for partner violence: are they based on ideology or evidence? *Legal and Criminological Psychology, 17,* 196-215. doi:10.1111/j.2044-8333.2011.02029

Dixon, L., Browne, K., & Hamilton-Giachritsis, C. (2005). Risk factors of parents abused as children: A mediational analysis of the intergenerational continuity of child maltreatment (Part I). *Journal of Child Psychology and Psychiatry, 46,* 47-57. doi: 10.1111/j.1469-7610.2004.00339.x

Dixon, L & Graham-Kevan, N. (2011). Understanding the nature of etiology of intimate partner violence and implications for practice and policy. *Clinical Psychology Review, 31,* 1145-1155. doi:10.1016/j.cpr.2011.07.001

Dobash, R. P., & Dobash, R. E. (2004). Women's violence to men in intimate relationships: Working on a puzzle. *British Journal of Criminology, 44,* 327-349. doi:10.1093/bjc/azh026

Dobash, R. E., Dobash, R. P., Wilson, M., & Daly, M. (1992). The myth of sexual symmetry in marital violence. *Social Problems, 39,* 71-91. doi:10.1525/sp.1992.39.1.03×00641

Dodge. K, Prince, J. & Newman, J. (1990). Hostile attribution biases in severely aggressive adolescents. *Journal of Abnormal Psychology, 99,* 385-392. doi:10.1037/0021-843x.99

Doumas, D. M., Pearson, C. L., Elgin, J. E. & McKinley, L. L. (2008). Adult attachment as a risk factor for intimate partner violence: The "mispairing" of partners' attachment styles. *Journal of Interpersonal Violence,* 23, 616-634. doi: 10.1177/0886260507313526

Dutton, D. G. (1995). Male abusiveness in intimate relationships. *Clinical Psychology Review, 15,* 567-581. doi:10.1016/0272-7358(95)00028-N

Dutton, D. G. (2006). Nested ecological theory. In D. G Dutton (Ed.), *Rethinking Domestic Violence* (pp. 18-40). Canada: UBC Press.

Dutton, D. G. (2007). *The abusive personality: Violence and control in intimate relationships (2nd ed.).* New York: Guilford Press.

Dutton, D. G. & Corvo, K. (2006) Transforming a flawed policy: A call to revive psychology and science in domestic violence research and practice. *Aggression and Violent Behavior, 11,* 457-483. doi:10.1016/j.avb.2006.01.007

Eckhardt, C. I., Murphy, C. M., Whitaker D. J., Sprunger, J., Dykstra, R., & Woodward, K. (2013). The effectiveness of intervention programs for perpetrators and victims of Intimate Partner Violence. *Partner Abuse, 4,* 196-231. doi:10.1891/1946-6560.42.196

Eckhardt, C. I., Norlander, B., & Deffenbacher, J. (2004). The assessment of anger and hostility: A critical review. *Aggression and Violent Behavior, 9,* 17-43. doi:10.1016/S1359-1789(02)001167

Eckhardt, C. I., Samper, R. E., & Murphy, C. M. (2008). Anger disturbances among perpetrators of intimate partner violence: Clinical characteristics and outcomes of court-mandated treatment. *Journal of Interpersonal Violence, 23,* 1600-1617. doi 10.1177/0886260508314322

Ehrensaft, M. K., Moffitt, T. E., & Caspi, A. (2004). Clinically abusive relationships in an unselected birth cohort: Men's and women's participation and developmental antecedents. *Journal of Abnormal Psychology, 133,* 258-271. Doi:10.1037/0021-843X.113.2.258

Eysenck, H. J., & Eysenck, S. B. G. (1975). *Manual of the Eysenck Personality Questionnaire.* London: Hodder and Stroughton, Ltd.

Farmer, E. & Callan, S. (2012). *Beyond violence: Breaking cycles of domestic abuse. A policy report for the Centre of Social Justice.* Available from http://www.centreforsocialjustice.org.uk/UserStorage/pdf/Pdf%20reports/DA%20Full%20report.pdf

Feder, L., Holditch Niolon, P., Campbell, J., Wallinder, J., Nelson, R., & Larrouy, H. (2011). The need for experimental methodology in intimate partner violence: Finding programs that effectively prevent IPV. *Violence Against Women, 17,* 340-358. doi:10.1177/1077801211398620

Feder, L., & Wilson, D. B. (2005). A meta-analytic review of court-mandated batterer intervention programs: Can courts affect abusers' behavior? *Journal of Experimental Criminology, 1,* 239-262. doi:10.1007/s11292-005-1179-0

Feiring, C., Deblinger, E., Hoch-Espada, A., & Haworth, T. (2002). Romantic relationship aggression and attitudes in high school students: The role of gender, grade, and attachment and emotional styles. *Journal of Youth and Adolescence, 31,* 373-385. doi:10.1023/A:1015680625391

Fellmeth, G. L. T., Heffernan, C., Nurse, J., Habibula, S., & Sethi, D. (2013). Educational and skills-based interventions for preventing relationship and dating violence in adolescents and young adults. *Cochrane Database of Systematic Reviews, 6,* 1-81. doi:10.1002/14651858.CD004534.pub3

Ferguson, C. J. (2009). An effect size primer: A guide for clinicians and researchers. *Professional Psychology: Research and Practice, 40,* 532-538. doi: 10.1037/a0015808.

Ferguson, C. J. & Dyck, D. (2012). Paradigm change in aggression research: The time has come to retire the General Aggression Model. *Aggression and Violent Behavior, 17,* 220-228. doi:10.1016/j.avb.2012.02.007

Fernandez-Montalvo, J., Echeburua, E., & Amor, P. J. (2005). Aggressors against women in prison and in the community: An exploratory study of a differential profile. *International Journal of Offender Therapy and Comparative Criminology, 49,* 158-167. doi:10.1177/0306624X04269005

Festinger, L. (1957). *A Theory of Cognitive Dissonance.* Stanford, CA: Stanford University Press.

Field, A. (2009). *Discovering Statistics using SPSS. Third Edition.* London: Sage Publications Ltd.

Field, A. (2013). *Discovering Statistics using SPSS. Fourth Edition.* London: Sage Publications Ltd.

Fisher, R. A. (1922). On the interpretation of chi square from the contingency tables, and the calculation of P. Journal of the Royal Statistical Society, 85, 87-94. doi:10.1214/ss/1177011454

Foley, P. F., Hartman, B. W., Dunn, A. B., Smith, J. E., & Goldberg, D. M. (2002). The utility of the State–Trait Anger Expression Inventory with offenders. *International Journal of Offender Therapy and Comparative Criminology, 46,* 364-378. doi:10.1177/0306624X02463009

Follingstad, D. R., Bradley, R. G., Helff, C. M., & Laughlin, J.E. (2002). A model for predicting dating violence: Anxious attachment, angry temperament, and need for relationship control. *Violence and Victims, 17,* 35-47. doi:10.1891/vivi.17.1.35.33639

Follingstad, D. R., Wright, S., Lloyd, S., & Sebastian, J. A. (1991). Sex differences in motivations and effects in dating violence. *Family Relations, 40,* 51-57.

Foshee, V. A., Bauman, K. E., Ennett, S. T., Suchindran, C., Benefield, T., & Linder, G. F. (2005). Assessing the effects of the dating violence prevention program "Safe Dates" using random coefficient regression modeling. *Prevention Science, 6,* 245-258. 10.1007/511121-005-0007-0

Foshee, V. A., Bauman, K. E., Linder, F., Rice, J., & Wilcher, R. (2007). Typologies of adolescent dating violence: Identifying typologies of adolescent dating violence perpetration. *Journal of Interpersonal Violence, 22,* 498-519. doi:10.1177/0886260506298829

Foshee, V. A., Linder, G. F., Bauman, K. E., Langwick, S. A., Arriaga, X. B., Heath, J. L., … Bangdiwala, S. (1996). The Safe Dates Project: theoretical basis, evaluation design, and selected baseline findings. *Journal of Preventive Medicine, 12,* 39-47.

Foshee, V. A., Linder, G. F., MacDougal, J. E., & Bangdiwala, S. (2001). Gender differences in the longitudinal predictors of adolescent dating violence. *Preventative Medicine, 32,* 128-141. doi:10.1006/pmed.2000.0793.

Foshee, V. A., McNaughton Reyes, H. L., Gottfredson, N. C., Chang, L. Y., Ennett, S. T. (2013). A longitudinal examination of psychological, behavioral, academic, and relationship consequences of dating abuse victimization among a primarily rural sample of adolescents. *Journal of Adolescent Health, 53,* 723-729. doi:10.1016/j.jadohealth.2013.06.016

Fraley, R. C, Waller, N. G., & Brennan, K. A. (2000). An Item Response Theory analysis of self-report measures of adult attachment. *Journal of Personality and Social Psychology, 28,* 350-365. doi:10.1037//0022-3514.78.2.350

Fritz, P. A. T., & Slep, A. M. S. (2009). Stability of physical and psychological adolescent dating aggression across time and partners. *Journal of Clinical Child and Adolescent Psychology, 38,* 303-314. doi:10.1080/15374410902851671.

Fuqua, D. R., Leonard, E., Masters, M. A., Smith, R. J., Campbell, J. L., & Fischer, P. C. (1991). A structural analysis of the State –Trait Anger Expression Inventory. *Educational and Psychological Measurement, 51,* 439-446. doi:10.1177/0013164491512018

Galambos, N. L., Petersen, A. C., Richards, M., & Gitelson, I. B. (1985). The Attitudes Towards Women Scale (AWSA): A study of reliability and validity. *Sex Roles, 13,* 343-356. doi:10.1007/BF00288090

Garson, G. D. (2012). *Life Tables and Kaplan-Meier Analysis: Nonparametric survival analysis (Blue Book Series).* Asheboro, NC: Statistical Associates Publishing.

Gentry, W. D., Chesney, A. P., Gary, H. G., Hall, R. P., & Harburg, E. (1982). Habitual anger-coping styles: I. Effect of mean blood pressure and risk for essential hypertension. *Psychosomatic Medicine, 44,* 195-202.

Gilbert, F., Daffern, M., Talevski, D., & Ogloff, J. R. P. (2013). The role of aggression-related cognition in the aggressive behavior of offenders: A General Aggression Model perspective. *Criminal Justice and Behavior, 40,* 119-138. doi:10.1177/0093854812467943

Gilchrist, E., Johnson, R., Takriti, R, Weston, S., Beech, A. R., & Kebbell, M. (2003). *Domestic violence offenders: characteristics and offending related needs.* Home Office Findings 217. Available from http://webarchive.nationalarchives.gov.uk/20110218135832/http:/rds.homeoffice.gov.uk/rds/pdfs2/r217.pdf

Glass, N., Fredland, N., Campbell, J., Yonas, M., Sharps, P. & Kub, J. (2003). Adolescent dating violence: Prevalence, risk factors, health outcomes, and implications for clinical practice. *Journal of Obstetric, Gynaecologic, & Neonatal Nursing, 32,* 227-238. doi:10.1177/0884217503252033

Golding, J. A. (1999). Intimate partner violence as a risk factor for mental disorders: A Meta-Analysis. *Journal of Family Violence, 14,* 99-132. doi:10.1023/A:1022079418229

Gondolf, E. W. (2001). Limitations of experimental evaluation of Batterer Programs. *Trauma Violence Abuse, 2,* 79-88. doi:10.1177/1524838001002001005

Gondolf, E.W. (2011). The weak evidence for batterer program alternatives. *Aggression and Violent Behavior, 16,* 347-353. doi:10.1016/j.avb.2011.04.011

Gordon, J. A., & Moriarty, L. J. (2003). The effects of domestic violence batterer treatment on domestic violence recidivism: The Chesterfield County experience. *Criminal Justice and Behaviour, 30,* 118-134. doi:10.1177/0093854802239166

Gupta, S. K. (2011). Intention-to-treat concept: A review. *Perspectives in Clinical Research, 2,* 109-112. doi:10.4103/2229-3485.83221

Hamel, J. (2007). Toward a gender-inclusive conception of intimate partner violence research and theory: Part 1 - traditional perspectives. *International Journal of Men's Health, 6,* 36-53. doi:10.3149/jmh.0601.36

Hamel, J. (2009). Toward a gender-inclusive conception of intimate partner violence research and theory: Part 2-New directions. *International Journal of Men's Health, 8,* 41-59. doi:10.3149/jmh.0801.41

Hamilton, L., Koehler, J. A., & Lösel, F. A. (2013). Domestic violence perpetrator programs in Europe, Part I: A survey of current practice. *International Journal of Offender Therapy and Comparative Criminology, 57,* 1189-1205. doi:10.1177/0306624X12469506

Hanson, R. K., Bourgon, G., Helmus, L., & Hodgson, S. (2009). *A meta-analysis of the effectiveness of treatment for sexual offenders: Risk, need, and responsivity.* Corrections Research, Public Safety Canada, Ottawa, Canada. Available from: http://www.publicsafety.gc.ca/res/cor/rep/2009-01-trt-so-eng.aspx

Harper, F. W. K., Austin, A. G., Cercone, J. J., & Arias, I. (2005). The role of shame, anger, and affect regulation in men's perpetration of psychological abuse in dating relationships. *Journal of Interpersonal Violence, 20,* 1648-1662. doi: 10.1177/0886260505278717

Haynie, D. L., Farhat, T., Brooks-Russell, A., Wang, J., Barbieri, B., & Iannotti, R. J. (2013). Dating violence perpetration and victimization among U.S. adolescents: Prevalence, patterns, and associations with health complaints and substance use. *Journal of Adolescent Health, 53,* 194-201. doi:10.1016/j.jadohealth.2013.02.008

Hester, M. (2006). Making it through the criminal justice system: Attrition and domestic violence. *Social Policy and Society, 5,* 79-90. doi:10.1017/S1474746405002769

Holtzworth-Munroe, A. (2000). A typology of men who are violent toward their female partners: Making sense of the heterogeneity in husband violence. C*urrent Directions in Psychological Science, 9,* 140-143. doi: 10.1111/1467-8721.00079

Holtzworth-Munroe, A., Rehman, U., & Herron, K. (2000). General and spouse-specific anger and hostility in subtypes of maritally violent men and nonviolent men. *Behavior Therapy, 31,* 603-630. doi:10.1016/S0005-7894(00)80034-9

Home Office. (2013). *Cross-Government Definition of Domestic Violence – A consultation summary of responses.* Available from https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/157800/domestic-violence-definition.pdf

Howard, P. (2006). *The Offender Assessment System: an evaluation of the second pilot.* Retrieved on 21st July, 2012, from http://webarchive.nationalarchives.gov.uk/20100623194820/rds.homeoffice.gov.uk/rds/pdfs06/r278.pdf

Hupka, R. B., & Rusch, P. A. (1989). Interpersonal Relationship Scale. In G. L. White & P. E. Mullen (Eds.), *Jealousy: Theory, research, and clinical strategies.* New York: The Guilford Press.

Hupka, R. B. & Rusch, P. A. (1992). Interpersonal Relationship Scale. In E.W. Mathes (Ed.), *Jealousy: The psychological data.* Lanham: University Press of America.

Hupka, R. B., & Rusch, P. A. (2001). Interpersonal Relationship Scale. In B. F. Perlmutter, J. Touliatos, & G. W. Holden (Eds.), *Handbook of family measurement techniques: Instruments and Index: Vol. 3. (pp. 148-150).* Thousand Oaks, CA: SAGE.

Jackson, S. M. (1999). Issues in the dating violence research: A review of the literature. *Aggression and Violent Behavior, 4,* 233-247. doi:10.1016/S1359-1789(97)00049-9

James, L. R., McIntyre, M. D., Glisson, C. A., Green, P. D., Patton, T. W., LeBreton, J. M., … Williams, L. J. (2005). Conditional reasoning: An efficient, indirect method for assessing implicit cognitive readiness to aggress. *Organizational Research Methods, 8,* 69-99. doi: 10.1177/1094428104272182

Johnson, M. P. (1995). Patriarchal terrorism and common couple violence: Two forms of violence against women. *Journal of Marriage and the Family, 57,* 283-294. doi:10.2307/353683

Johnson, M. P. (2006). Conflict and control: Gender, symmetry, and asymmetry in domestic violence. *Violence Against Women, 12,* 1003-1018. doi: 10.1177/1077801206293328

Johnson, M. P. (2008). *A typology of domestic violence: Intimate terrorism, violent resistance, and situational couple violence.* Boston, MA: Northeastern University Press.

Kaplan, R. M., & Saccuzzo, D. P. (2009). *Psychological Testing: Principles, applications, and issues (7th ed.)* Belmont, CA: Cengage Learning.

Khodayarifard, M., Spielberger, C. D., Lavasani, M. G., & Zardkhaneh, S. A. (2013). Psychometric properties of Farsi version of the Spielberger's State-Trait Anger Expression Inventory-2 (FSTAXI-2). *Social and Behavioral Sciences, 82,* 325-329. doi: 10.1016/j.sbspro.2013.06.269.

Kline, P (1993). *The handbook of psychological testing.* London: Routledge.

Koss, M. P., Gidycz., C. A., & Wisniewski, N. (1987). The scope of rape: Incidence and prevalence of sexual aggression and victimisation in a national sample of higher education students. *Journal of Consulting and Clinical Psychology, 55,* 162-170.

Kroner, D. G., & Reddon, J. R. (1992). The anger expression scale and state-trait anger scale: Stability, reliability and factor structure in an inmate sample. *Criminal Justice and Behavior, 19,* 397-408. doi:10.1177/0093854892019004004

Kroner, D., Reddon, J. R., & Serin, R. C. (1992). The Multidimensional Anger Inventory: Reliability and factor structure in an inmate sample. *Educational and Psychological Measurement, 52,* 687-693. doi:10.1177/0013164492052003018

Kropp, P. R., Hart, S. D., Webster, C. D., & Eaves, D. (1995). *Manual for the Spousal Assault Risk Assessment guide (2nd ed.).* Vancouver, BC: British Columbia Institute Against Family Violence

Lawrence, E., Orengo-Aguayo, R., Langer, A. & Brock, R. L. (2012). The impact and consequences of partner abuse on partners. *Partner Abuse, 3,* 406-428. doi:10.1891/1946-6560.3.4.406

Laws, D. R., & Ward, T. (2011). *Desistance and sexual offending: Alternatives to throwing away the keys*. New York, NY: Guilford Press.

Leen, E., Sorbring, E., Mawer, M., Holdsworth, E., Helsing., B., & Bowen, E. (2013). Prevalence, dynamic risk factors and the efficacy of primary interventions for adolescent dating violence: An international review. *Aggression and Violent Behavior 18,* 159-174. doi:10.1016/j.avb.2012.11.015

Levene, H. (1960). Robust tests for equality of variances. In I. Olkin (Ed.). *Contributions to probability and statistics*. Palo Alto, CA: Stanford University Press.

Lindqvist, J. K., Daderman, A. M., & Hellstrom, A. (2003). Swedish adaptations of the Novaco Anger Scale -1998, The Provocation Inventory and the State-Trait Anger Expression Inventory 2. *Social Behaviour and Personality, 31,* 773-788. doi: 10.2224/sbp.2003.31.8.773

Loeber, R. & Farrington, D. P. (2014). Age Crime Curve. In G. Bruinsma & D. Weisburd (Eds.). *Encyclopedia of Criminology and Criminal Justice.* (pp. 12-18). New York, NY: Springer.

Loeber, R., Menting, B., Lynam, D. R., Moffitt, T. E., Stouthamer-Loeber, M., Stallings, R., Farrington, D. P., & Pardini, D. (2012). Findings from the Pittsburgh Youth Study: Cognitive impulsivity and intelligence as predictors of the Age–Crime Curve, *Journal of the American Academy of Child and Adolescent Psychiatry, 51,* 1136 -1149. doi: 10.1016/j.jaac.2012.08.019Lösel, F. & Schmuker, M. (2005). The effectiveness of treatment for sexual offenders: A comprehensive meta-analysis. *Journal of Experimental Criminology, 1,* 117-146. doi:10.1007/s11292-004-6466-7

Loza, W., & Loza-Fanous, A. (1999a). The fallacy of reducing rape and violent recidivism by treating anger. *International Journal of Offender Therapy and Criminology, 43,* 492-502. doi:10.1177/0306624X99434007

Loza, W., & Loza-Fanous, A. (1999b). Anger and prediction of violent and nonviolent offenders' recidivism. *Journal of Interpersonal Violence, 14,* 1014-1029. doi:10.1177/088626099014010002

Lussier, P., Farrington, D. P., & Moffitt, T. E. (2009). Is this antisocial child father of the abusive man? A 40 year prospective longitudinal study on the developmental antecedents of intimate partner violence. *Criminology, 47,* 741-780. doi:10.1111/j.1745-9125.2009.00160.x

Mahmood, S. (2012). *Introduction to survival analysis, Kaplan Meier estimate.* Available from http://www.academia.edu/1713127/Introduction_to_survival_analysis_Kaplan_Meier_estimate

Maltz, M. D. (2001). *Recidivism*. Orlando, FL; Academic Press, Inc. Available from http://www.uic.edu/depts/lib/forr/pdf/crimjust/recidivism.pdf

Mann, H. B., & Whitney, D. R. (1947). On a test of whether one of two random variables is stochastically larger than the other. *Annals of Mathematical Statistics, 18,* 50-60. doi:10.1214/aoms/1177730491

Margolin, G., & Foo, L. (1992). *The Attitudes about Dating Scale.* (Unpublished measure). University of Southern California, California.

Maxwell, J. P., Sukhodolsky, D. G., & Sit, C. H. P. (2009). Preliminary validation of a Chinese version of the State-Trait Anger Expression Inventory-2. *Asian Journal of Social Psychology, 12,* 1-11. doi: 10.1111/j.1467-839X.2008.01264.x

McGrath, R. E. & Meyer, G. J. (2006). When effect sizes disagree: The case of r and d. *Psychological Methods, 11,* 386-401. doi:10.1037/1082-989X.11.4.386

McGuire, J. (2013).What Works to reduce reoffending: 18 years on. In L. A. Craig, L. Dixon & T. A. Gannon (Eds.). *What Works in Offender Rehabilitation: An evidence based approach to assessment and treatment* (pp. 20-49). Chichester: Wiley-Blackwell. doi:10.1002/9781118320655

McMurran, M., & Theodosi, E. (2007). Is treatment non-completion associated with increased reconviction over no treatment? *Psychology, Crime and Law, 13,* 333-343. doi:10.1080/10683160601060374

Mela, M., Balbuena, L., Randy, D. C., Wong, S., Gu, D., Polvi, N., & Gordon, A. (2008). The STAXI as a measure of inmate anger and a predictor of institutional offending. *Journal of Forensic Psychiatry & Psychology, 19,* 396-406. doi:10.1080/14789940802164090

Miguel-Tobal, J. J., Casado, M. I., Cano-Vindel, A., & Spielberger, C. D. (2001). *Inventario de Expresión de Ira Estado- Rasgo–STAXI.* Madrid: TEA Ediciones.

Miller, W. R., & Rollnick, S. (2002). *Motivational Interviewing: Preparing People For Change.* New York: Guilford.

Millon, T., Millon, C., Davis, R., & Grossman, S. (2006). *Millon Clinical Multiaxial Inventory: MCMI-III Manual.* Minneapolis, MN: Pearson Education, Inc

Mills, J. F., & Kroner, D. G. (2003). Anger as a predictor of institutional misconduct and recidivism in a sample of violent offenders. *Journal of Interpersonal Violence, 18,* 282-294. doi:10.1177/0886260502250085

Ministry of Justice (2010). *What works with Domestic Violence offenders?* Available from http://hlsweb.dmu.ac.uk/pgcpd/roh436/official-documents/What_works_-_Domestic_violence[1].pdf.

Medeiros, R. A., & Straus, M. A. (2006b). Risk factors for physical violence between dating partners: Implications for gender-inclusive prevention and treatment of family violence. In J. C. Hamel & T. Nicholls (Eds.). *Family Approaches to Domestic Violence*. New York: Springer.

Moffitt, T. E., Caspi, A., Rutter, M., & Silva, P. A. (Eds.) (2001). *Sex differences in antisocial behaviour: Conduct disorder, delinquency, and violence in the Dunedin Longitudinal Study.* Cambridge: Cambridge University Press.

Munoz-Rivas, M. J., O'Leary, K. D., & Gonzales, M. P. (2007). Aggression in adolescent dating relationships: Prevalence, justification, and health consequences. *Journal of Adolescent Health, 40,* 298-304. doi:10.1016/j.jadohealth.2006.11.137

Murphy, C. M., Taft, C. T., & Eckhardt, C. I. (2007). Anger problem profiles among partner violent men: Differences in clinical presentation and treatment outcome. *Journal of Counselling Psychology, 54,* 189-200. doi:10.1037/0022-0167.54.2.189

Narayan, A. J., Englund, M. M., Carlson, E. A. & Egeland, B. (2014). Adolescent conflict as a developmental process in the prospective pathway from exposure to interparental violence to dating violence. *Journal of Abnormal Child Psychology, 42,* 239-250. doi:10.1007/s10802-013-9782-4

NOMS Test Battery Guide (2012). *IDAP/CDVP Test Battery Guide. February 2012.* Ministry of Justice.

Norlander, B., & Eckhardt. C. (2005). Anger, hostility, and male perpetrators of intimate partner violence: A meta-analytic review. *Clinical Psychology Review, 25,* 119-152. doi:10.1016/j.cpr.2004.10.001

Novaco, R.W. (1994). *Novaco Anger Scale and Provocation Inventory (NAS-PI).* Los Angeles, CA: Western Psychological Services.

Novaco, R.W., Ramm, M., & Black, L. (2000). Anger treatment with offenders. In C. R. Hollin (Ed), *The essential handbook of offender assessment and treatment* (pp. 281-296). London: John Wiley and Sons.

O'Connor, D. B., Archer, J., & Wu, F. W. C. (2001). Measuring aggression: Self-reports, partner reports, and responses to provoking scenarios. *Aggressive Behavior, 27,* 79-101. doi:10.1002/ab.2

Office for National Statistics. (2013). *Statistical Bulletin. Focus on: Violent Crime and Sexual Offences, 2011/12.* Available from http://www.ons.gov.uk/ons/dcp171778_298904.pdf

O'Keefe, M. (1997). Predictors of dating violence among high school students. *Journal of Interpersonal Violence, 12,* 546-569. doi:10.1177/088626097012004005

O'Keefe, M. (2005). Teen Dating Violence: A review of risk factors and prevention efforts. *National Electronic Network on Violence Against Women.* Available from http://www.unajauladeoro.com/cd/documentos/AR_TeenDatingViolence.pdf

O'Leary, D. K., Barling, J., Arias, I., Rosenbaum, A., Malone, J., & Tyree, A. (1989). Prevalence and stability of physical aggression between spouses: A longitudinal analysis. *Journal of Consulting and Clinical Psychology, 57,* 263-268. doi:10.1037/0022-006X.57.2.263

O'Leary, K. D., Smith-Slep, A.M., & O'Leary, S.G. (2007). Multivariate models of men's and women's partner aggression. *Journal of Consulting and Clinical Psychology, 75,* 752-764. doi: 10.1037/0022-006X.75.5.752

Olver, M. E., Stockdale, K. C., & Wormith, J. S. (2011). A meta-analysis of predictors of offender treatment dropout and its relationship to recidivism. *Journal of Consulting and Clinical Psychology, 79,* 6-21. doi: 10.1037/a0022200

Pearson, K. (1900). On the criterion that a given system of deviations from the probably in the case of a correlated system of variables is such that it can be reasonably supposed to have arisen from random sampling. *Philosophical Magazine, 50,* 157-175.

Pence, E., & Paymar, M. (1993). *Education groups for men who batter: The Duluth model.* New York: Springer.

Price, E. L., Byers, E. S., Belliveau, N., Bonner, R., Caron, B., Doiron, D., et al. (1999). The Attitudes Towards Dating Violence Scales: Development and initial validation. *Journal of Family Violence, 14,* 351-375. doi:10.1023/A:1022830114772

Prospero, M. (2006). The role of perceptions in dating violence among young adolescents. *Journal of Interpersonal Violence, 21,* 470-484. doi:10.1177/088620505285725

Puente, S., & Cohen, D. (2003). Jealousy and the meaning (or non-meaning) of violence. *Personality and Social Psychology Bulletin, 29,* 449-460. doi:10.1177/0146167202250912

Rea, L. M., & Parker, R. A. (1992). *Designing and conducting survey research.* San Francisco: Jossey-Boss.

Reeves, P. M., & Orpinas, P. (2012). Dating norms and dating violence among ninth graders in Northeast Georgia: Reports from student surveys and focus groups. *Journal of Interpersonal Violence, 27,* 1677-1698. doi:10.1177/0886260511430386

Respect. (2008). *Respect position statement: gender and domestic violence.* Available from http://www.respect.uk.net/pages/briefing-papers.html.

Rosenfeld, B. D. (1992). Court ordered treatment of spouse abuse. *Clinical Psychology Review, 12,* 205–226. doi:10.1016/0272-7358(92)90115-O

Saunders, D. G. (2008). Group interventions for men who batter: A summary of program descriptions and research. *Violence and Victims, 23,* 156-172. doi:10.1891/0886-6708.23.2.156

Schultz, S. D. (1954). A differentiation of several forms of hostility by scales empirically constructed from significant items of the Minnesota Personality Inventory. (*Unpublished Doctoral Thesis).* Pennsylvania State College, Pennsylvania.

Schumacher, J. A., & Smith Slep, A. M. (2004). Attitudes and dating aggression: A cognitive dissonance approach. *Prevention Science, 5,* 231-243. doi:10.1023/B:PREV.0000045357.19100.77

Schwartz, M., O'Leary, S. G. & Kendziora, K. T. (1997). Dating aggression among high school students. *Violence and Victims, 12,* 295-305.

Scott, K., & Straus, M. (2007). Denial, minimization, partner blaming, and intimate aggression in dating partners. *Journal of Interpersonal Violence, 22,* 851-871. doi:10.1177/0886260507301227

Sears, H. A., Byers, E. S., & Price, L. (2007). The co-occurrence of adolescent boys' and girls' use of psychologically, physically, and sexually abusive behaviours in their dating relationships. *Journal of Adolescence, 30*, 487-504. doi:10.1016/j.adolescence.2006.05.002

Shorey, R. C., Brasfield, H., Febres, J., & Stuart, G. L. (2011). The association between impulsivity, trait anger, and the perpetration of intimate partner and general violence among women arrested for domestic violence. *Journal of Interpersonal Violence, 26*, 2681-2697 doi: 10.1177/0886260510388289

Shorey, R. C., Cornelius, T. L., & Bell, K. M. (2008). A critical review of theoretical frameworks for dating violence: Comparing the dating and marital fields. *Aggression and Violent Behavior, 13,* 185-194. doi:10.1016/j.avb.2008.03.003

Shorey, R. C., Zucosky, H., Brasfield, H., Febres, J., Cornelius, T. L., Sage, C., & Stuart, G. L. (2012). Dating violence prevention programming: Directions for future interventions. *Aggression and Violent Behavior, 17,* 289-296. doi:10.1016/j.avb.2012.03.001

Shortt, J. W., Capaldi, D. M., Kim, H. K., Kerr, D. C. R., Owen, L. D., & Feingold, A. (2012). Stability of intimate partner violence by men across 12 years in young adulthood: Effects of relationship transitions. *Prevention Science, 13,* 360-369. doi:10.1007/s11121-011-0202-0

Sibley, C. G., & Liu, J. H. (2004). Short-term temporal stability and factor structure of the revised experiences in close relationships (ECR-R) measure of adult attachment. *Personality and Individual Differences, 36,* 969-975.

Siegel, J. M. (1986). The Multidimensional Anger Inventory (MAI). *Journal of Personality and Social Psychology, 51,* 191-200. doi:10.1037/0022-3514.51.1.191

Simon, T. R., Miller, S., Gorman-Smith, D., Orpinas, P., & Sullivan, T. (2010). Physical dating violence norms and sixth-grade students from four US Sites. *Journal of Early Adolescence, 30,* 395-409. doi:10.1177/0272431609333301

Spielberger, C. D. (1988). *Manual for the State-Trait Anger Expression Inventory,* Odessa, FL: Psychological Assessment Resources Inc.

Spielberger, C. D. (1999). *The State-Trait Anger Expression Inventory-2 (STAXI-2): Professional manual*. Odessa, FL: Psychological Assessment Resources Inc.

Spielberger, C. D. (2006). Cross-cultural assessment of emotional states and personality traits. *European Psychologist 11,* 297-303. doi:10.1027/1016-9040.11.4.297

Spielberger, C. D., Krasner, S. S., & Solomon, E. P. (1988). The experience, expression, and control of anger. In M. P. Janisse (Ed.), *Health psychology: Individual differences and stress* (pp. 89–108). New York, NY: Springer.

Stewart, L., Gabora, N., Kropp, R., & Lee, Z. (2005). *Family violence programming: Treatment outcome for Canadian federally sentenced Offenders.* Available from http://www.csc-scc.gc.ca/research/r174-eng.shtml#Toc183336165.

Stith, S. M., McCollum, E. E., Amanor-Boadu, Y., & Smith, D. (2012). Systemic perspectives on Intimate Partner Violence treatment. *Journal of Marital and Family Therapy, 38,* 220-240. doi: 10.1111/j.1752-0606.2011.00245.x

Stith, S. M., Rosen, K. H., Middleton, K. A., Busch, A. L. Lundeberg, K., & Carlton, R. P. (2000). The intergenerational transmission of Spouse Abuse: A Meta-Analysis. *Journal of Marriage and Family, 62,* 640-654.

Stith, S. M., Smith, D. B., Penn, C. E., Ward, D. B., & Tritt, D. (2004). Intimate partner physical abuse perpetration and victimization risk factors: A meta-analytic review. *Aggression and violent behaviour, 10,* 65-98.

Stover, C. S., Meadows, A. L., & Kaufman, J. (2009). Interventions for intimate partner violence: Review and implications for evidence-based practice. *Professional Psychology: Research and Practice, 40, 223-233*. doi: 10.1037/a0012718

Straus, M. A. (1979). Measuring intra-familial conflict and violence: The Conflict Tactics Scales (CTS). *Journal of Marriage and the Family, 41*, 75-88.

Straus, M. A. (2008). Dominance and symmetry in partner violence by male and female university students in 32 nations. *Children and Youth Services Review, 30,* 252-275. doi:10.1016/j.childyouth.2007.10.004

Straus, M. A. (2011). Gender symmetry and mutuality in perpetration of clinical-level partner violence: Empirical evidence and implications for prevention and treatment. *Aggression and Violent Behavior, 16,* 279-288. doi:10.1016/j.avb.2011.04.010

Strickland, P. (2013). *Library Standard Note SN/HA/6337, Domestic Violence*. Home Office Section.

Swaffer, T. & Epps, K. (1999). The psychometric assessment of anger in male and female adolescents' resident at a secure youth treatment centre. *Journal of Adolescence, 22,* 419-422. doi:10.1006/jado.1999.0233

Swahn, M. H., Simon, T., Arias, I., & Bossarte, R. M. (2008). Measuring sex differences in violence victimization and perpetration within date and same-sex peer relationships. *Journal of Interpersonal Violence, 23,* 1120-1138. doi:10.1177/0886260508314086

Taft, C. T., Murphy, C. M., King, D. M., Musser, P. H., & DeDeyn, J. M (2003). Process and treatment adherence factors in group cognitive-behavioural therapy for partner violent men. *Journal of Consulting and Clinical Psychology, 71,* 812-820.

Tangney, J. P., Wagner, P. E., Gavlas, J., & Garmzow, R. (1991). *The Test of The Self-Conscious Affect for Adolescents (TOSCA-A).* Fairfax, VA: George Mason University.

Temple, J. R., Shorey, R. C., Tortolero, S. R., Wolfe, D. A., & Stuart, G. L. (2013). Importance of gender and attitudes about violence in the relationship between exposure to interparental violence and the perpetration of teen dating violence. *Child Abuse and Neglect, 37,* 343-352. doi:10.1016/j.chiabu.2013.02.001

Theobald, D., & Farrington, D. P. (2014). Child and adolescent predictors of male intimate partner violence. *Journal of Child Psychology and Psychiatry 53,* 1242-1249. doi:10.1111/j.1469-7610.2012.02577.x

Thornton, A. J., Graham-Kevan, N., & Archer, J. (2012). Prevalence of women's violent and non-violent offending behaviour: A comparison of self-reports, victims' reports and third-party reports. *Journal of Interpersonal Violence, 27,* 1399- 1427. doi:10.1177/0886260511425789

Tjarden, P., & Thoennes, N. (2000). Prevalence and consequences of male-to-female and female-to-male intimate partner violence as measured by the National Violence Against Women Survey. *Violence against Women, 6,* 142-161. doi: 10.1177/1077801002218176

Vagi, K. J., Rothman, E., Latzman, N. E., Teten Tharp, A., Hall, D. M., & Breiding, M. (2013). Beyond correlates: A review of risk and protective factors for adolescent dating violence perpetration. *Journal of Youth and Adolescence, 42,* 633-649. doi:10.1007/s10964-013-9907-7

Varley Thornton, A. J., Graham-Kevan, N., & Archer, J. (2010). Adaptive and maladaptive personality traits as predictor of violent and nonviolent offending behaviour in men and women. *Aggressive Behavior, 36,* 177-186. doi: 10.1002/ab.20340

Walker, K., Bowen, E. & Brown, S. (2013). Desistance from intimate partner violence: A critical review. *Aggressive and Violent Behavior, 18,* 271-280. doi:10.1016/j.avb.2012.11.019

Ward, T., & Maruna, S. (2007). *Rehabilitation: Beyond the risk paradigm.* Oxon: Routledge.

Ward, T., Melser, J., & Yates, M. (2007). Reconstructing the Risk-Need-Responsivity model: A theoretical elaboration and evaluation. *Aggression and Violent Behavior, 12,* 208-228. doi:10.1016/j.avb.2006.07.001

Ward, T., Rose. C., & Willis, G. M. (2011). Offender Rehabilitation: Good Lives, desistance and risk reduction. In G. Davies & A. R. Beech (Eds.). *Forensic psychology: Crime, justice, law, interventions (2nd ed.).* (pp. 408-423). Chichester: John Wiley & Sons.

Ward, T., & Stewart, C. (2003). Criminogenic needs and human needs: A theoretical model. *Psychology, Crime & Law, 9,* 125-143. doi: 10.1080/1068316031000116247

Weatherstone, P (n.d). *New Directions in NOMS Domestic Abuse Programmes.* Ministry of Justice/ National Offender Management Service. Accessed from http://www.cepprobation.org/uploaded_files/Pres%20STARR%20Par%20Weatherstone.pdf

Wekerle, C. & Wolfe, D. A. (1999). Dating violence in mid-adolescence: Theory, significance, and emerging prevention initiatives. *Clinical Psychology Review, 19,* 435-456. doi:10.1016/S0272-7358(98)00091-9

Wilcoxon, F. (1945). Individual comparisons by ranking methods. *Biometrics Bulletin, 1,* 80-83. doi:10.2307/3001968

Wilke, D. J. & Vinton, L. (2005). The nature and impact of Domestic Violence across age cohorts. *Affilia, 20,* 316 - 328. doi:10.1177/0886109905277751

Williams, T. S., Connolly, J., Pepler, D., Craig, W., & Laporte, L. (2008). Risk models of dating aggression across different adolescent relationships: A developmental

psychopathology approach. *Journal of Consulting and Clinical Psychology, 76,* 622-632. doi: 10.1037/0022-006X.76.4.622

Wolf, K. A. & Foshee, V. A. (2003). Family violence, anger expression styles, and adolescent dating violence. *Journal of Family Violence, 18,* 309-316. doi: 10.1023/A:1026237914406

Wolfe, D. A., Wekerle, C., Reitzel-Jaffe, D., & Lefebvre, L. (1998). Factors associated with abusive relationships among maltreated and non-maltreated youth. *Development and Psychopathology, 10,* 61-85. doi:10.1017/S0954579498001345

Zakzanis, K. K. (2001). Statistics to tell the truth, the whole truth, and nothing but the truth: formulae, illustrative numerical examples, and heuristic interpretation of effect size analyses for neuropsychological researchers. *Archives of Clinical Neuropsychology, 16,* 653-667. doi:S0887-6177(00)00076-

Zweig, J. M., Yahner, J., Dank., M. & Lachman, P. (2014). Can Johnson's Typology of Adult Partner Violence apply to Teen Dating Violence? *Journal of Marriage and Family, 76,* 808-825. doi: 10.1111/jomf.12121

Appendices

Appendix One: Scoping search terms and syntax

Cochrane Database of Systematic Reviews (CDSR) and DARE database search terms

"domestic violence" or "intimate partner violence" or "dating violence" or "interpersonal violence" or "relationship violence" or "domestic abuse" or "interpersonal abuse" or "dating abuse" or "relationship abuse" or "intimate partner abuse" or "dating relationship violence" or "dating relationship abuse" or "dating relationship aggression" or "dating relationship" or "teenage partner abuse" or "teenage partner violence" or "teen dating violence" or "teen dating abuse"

AND: adolescen* or teen* or "young person" or boy* or girl* or "young male" or "young female*" or juvenil* or "high school student*" or student*

AND: attitude* or "pro-violent attitude*" or perception* or belief* or norm* or justif* or accept*

AND: offend* or perpetrat*

Scoping search terms and syntax in PsycINFO

	Terms Used	Number of Results
1	("domestic violence" or "intimate partner violence" or "dating violence" or "interpersonal violence" or "relationship violence" or "domestic abuse" or "interpersonal abuse" or "dating abuse" or "relationship abuse" or "intimate partner abuse" or "dating relationship violence" or "dating relationship abuse" or "dating relationship aggression" or "dating relationship*" or "teenage partner abuse" or "teenage partner violence" or "teen dating abuse" or "teen dating violence").mp. [mp=title, abstract, heading word, table of contents, key concepts, original title, tests & measures]	17713

2	(adolescen* or teen * or "young person" or boy* or girl* or "young male*" or "young female*" or juvenil* or "young offender" or "high school student*" or student*).mp. [mp=title, abstract, heading word, table of contents, key concepts, original title, tests & measures]	723524
3	(attitude* or “pro-violent attitude*” or perception* or belief* or norm* or justif* or accept*).mp. [mp=title, abstract, heading word, table of contents, key concepts, original title, tests & measures]	951993
4	("Systematic Literature Review" or "Literature Review" or "meta analys*" or review).mp. [mp=title, abstract, heading word, table of contents, key concepts, original title, tests & measures]	302335
5	1 2 3 and 4	79

Appendix Two: Main literature search terms and syntax for each electronic database.

PsycINFO 1987 to 2014, May, Week 2.

	Terms Used	Number of Results
1	("domestic violence" or "intimate partner violence" or "dating violence" or "interpersonal violence" or "relationship violence" or "domestic abuse" or "interpersonal abuse" or "dating abuse" or "relationship abuse" or "intimate partner abuse" or "dating relationship violence" or "dating relationship abuse" or "dating relationship aggression" or "dating relationship*" or "teenage partner abuse" or "teenage partner violence" or "teen dating abuse" or "teen dating violence").mp. [mp=title, abstract, heading word, table of contents, key concepts, original title, tests & measures]	17022
2	(adolescen* or teen * or "young person" or boy* or girl* or "young male*" or "young female*" or juvenil* or "young offender" or "high school student*" or student*).mp. [mp=title, abstract, heading word, table of contents, key concepts, original title, tests & measures]	517736
3	(attitud* or perception* or belief* or norm* or justif* or “acceptance of violence” or “pro violent attitudes”).mp. [mp=title, abstract, heading word, table of contents, key concepts, original title, tests & measures]	707078
4	(perpetrat* or victimis* or outcome* or prevalence or frequenc* or "base rate*").mp. [mp=title, abstract, heading word, table of contents, key concepts, original title, tests & measures]	399765
5	1 and 2 and 3 and 4	463
6	limit 5 to english language	436
7	limit 6 last 20 years	418

EMBASE 1988 to 2014 Week 20

	Terms Used	Number of Results
1	("domestic violence" or "intimate partner violence" or "dating violence" or "interpersonal violence" or "relationship violence" or "domestic abuse" or "interpersonal abuse" or "dating abuse" or "relationship abuse" or "intimate partner abuse" or "dating relationship violence" or "dating relationship abuse" or "dating relationship aggression" or "dating relationship*" or "teenage partner abuse" or "teenage partner violence" or "teen dating abuse" or "teen dating violence").mp. [mp=title, abstract, heading word, table of contents, key concepts, original title, tests & measures]	12321
2	(adolescen* or teen* or "young person" or boy* or girl* or "young male*" or "young female*" or juvenil* or "young offender" or "high school student*" or student*).mp. [mp=title, abstract, heading word, table of contents, key concepts, original title, tests & measures]	1359699
3	(attitud* or perception* or belief* or norm* or justif* or "acceptance of violence" or "pro violent attitudes").mp. [mp=title, abstract, heading word, table of contents, key concepts, original title, tests & measures]	2683664
4	(perpetrat* or victimis* or outcome* or prevalence or frequenc* or "base rate*").mp. [mp=title, abstract, heading word, table of contents, key concepts, original title, tests & measures]	2820997
5	1 and 2 and 3 and 4	423
6	limit 5 to english language	405
7	limit 6 to last 20 years	401

MEDLINE 1946 to 2014 May Week 1

	Terms Used	Number of Results
1	("domestic violence" or "intimate partner violence" or "dating violence" or "interpersonal violence" or "relationship violence" or "domestic abuse" or "interpersonal abuse" or "dating abuse" or "relationship abuse" or "intimate partner abuse" or "dating relationship violence" or "dating relationship abuse" or "dating relationship aggression" or "dating relationship*" or "teenage partner abuse" or "teenage partner violence" or "teen dating abuse" or "teen dating violence").mp. [mp=title, abstract, original title, name of substance word, subject heading word, keyword heading	9983

	word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier]	
2	(adolescen* or teen* or "young person" or boy* or girl* or "young male*" or "young female*" or juvenil* or "young offender" or "high school student*" or student*).mp. [mp=title, abstract, original title, name of substance word, subject heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier]	1905092
3	(attitud* or perception* or belief* or norm* or justif* or “acceptance of violence” or “pro violent attitudes”).mp. [mp=title, abstract, original title, name of substance word, subject heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier]	2091327
4	(perpetrat* or victimis* or outcome* or prevalence or frequenc* or "base rate*").mp. [mp=title, abstract, original title, name of substance word, subject heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier]	2238749
5	1 and 2 and 3 and 4	324
6	limit 5 to english language	314
7	limit 6 to last 20 years	309

Appendix Three: Inclusion/Exclusion Criteria

Inclusion/Exclusion Criteria	Criteria Met?	Any comments
Population Does the population consist of adolescents (male and female or male only) aged between 11 & 21 years of age? Is the population from UK or USA? Does the population consist of participants in the community?	Yes No Unclear Yes No Unclear Yes No Unclear	
Exposure: Has attitudes towards dating violence been assessed?	Yes No Unclear	
Outcomes: Has incidents of dating violence (perpetration and victimisation) been measured?	Yes No Unclear	
Study Design Cross Sectional/ Cohort Quantitative	Yes No Unclear	
Language Is the first language of the study English?	Yes No Unclear	

Studies that answer Yes to all questions to be included in the Review.

Appendix Four: Data extraction form

General Information

Date of Data Extraction

Title of Publication

Author (s)

Journal Title

Name of Reviewer

General Notes:

Verification of study meeting inclusion/exclusion criteria

Population:	Adolescents aged between 11 & 21 years old. UK or US population	Y	N	?
	Community Population (excluding offender/ psychiatric populations)	Y	N	?
Intervention	Exposure to questionnaire/ assessment exploring perceptions/attitudes towards dating violence	Y	N	?
Outcomes	Reported experience (perpetration or victimisation) of dating violence	Y	N	?
Language	English Only			

Study Design	Cross Sectional	Cohort	Other
Study to be carried forward?	Yes	No	

Specific Information of study

Methodological quality of the study

1. Study Design
2. Recruitment procedures
3. Quality assessment

Population:

1) Target Population

2) Inclusion Criteria

3) Exclusion Criteria

4) Recruitment procedures used

Number of Particpants:

Male: Female:

Number of participants refused:

Reasons:

Age of participants: Average of participants:

Ethnicity:

Number of participants reported to be currently in a relationship at time of study? (if possible)

Parental Consent Obtained?

Number of participants refused/ excluded following obtaining parental consent?

Other information:

How were participants recruited?

Sampling type?

Exposure:

- Use of a psychometric/ questionnaire focussed upon measuring attitudes/ perceptions of Dating Violence?
- Which psychometric/ questionnaires were used?
- Validated?
- Computer/ by hand for completion?
- What were the psychometrics measuring specifically?
- Mediating variables?

- Was the exposure conducted in a suitable/ confidential environment?
- Who administered the assessment?

Notes:

Outcome:

- Assessments administered to measure prevalence of Dating Violence?
- What assessment was used?
- Validated type of assessment?
- Did it rely on self-report?

What was measured?

Any drop outs?

Analysis

1) Statistic techniques Used

2) Direction and significance of results?

2) Were confounding variables assessed?

3) Were attrition rates adjusted for in statistics?

4) Were the statistics and results reported clearly?

5) Any over/ under stated significance in results?

Notes:

Appendix Five: Quality Assessment Form

Title of Study_____

Authors_____

Question	Yes (2)	Partial (1)	No (0)	Unclear (0)	Comments
Initial Screening					
Are the hypothesis clearly stated?					
Does the study include assessment of attitudes among adolescents towards dating violence within it?					
Design					
Has the study addressed the questions it has outlined?					
Is the design an appropriate way of answering the question?					
Selection Bias					
Is the sample representative of the population to which the results will generalise?					
Was a sufficient sample size used?					
Was the sample size justified?					
Were characteristics of the population clearly described?					
Were the independent and dependent variables clearly stated?					
Was the potential confounding variables controlled for or acknowledged?					
Measurement/Detection Bias					
Has incidents of dating violence been clearly defined and measured?					
Have the assessments used been clearly described, measured and standardised?					
Were participants blind?					
Were self-report measures used?					
Did all participants complete all /outcome measures?					
Were outcome measures validated?					
Attrition Bias					

Were reasons explained for those refusing to participate in the study?					
Have “drop outs” introduced high level of bias? (Reverse scored)					
Outcome Bias					
Was outcome measured in a correct way?					
Were the measures valid and reliable for the population?					
Statistics					
Was the statistical analysis used correctly?					
Were there attempts to deal with missing data?					
Result Bias					
Are the results logical?					
Have the results been clearly reported?					
Have all results been accounted for?					
Have limitations been discussed?					
Generalisability of findings					
Are the participants representative of the population studied?					
Can results be generalised to a wider population?					
Are the results consistent with other findings in available literature?					
COLUMN TOTAL					
OVERALL TOTAL					
Percentage					

Appendix Six: Participant consent form

COMMUNITY DOMESTIC VIOLENCE PROGRAMME

STATEMENT OF UNDERSTANDING

Probation staff have explained the programme and given me a leaflet.

I agree to attend the programme on these conditions:

1. Attendance

- I must attend all group sessions and regular one-to-one sessions arranged by my Case Manager.
- I must be on time. If I am late I may not be allowed to join the session and this will count as a session I have missed.
- If I am excluded from a session because of my behaviour, this will also count as a missed session.
- I must attend all sessions of the first module of the programme. If I miss any of them I must restart the programme.
- I cannot do more than one catch-up session in any of the other modules. If I miss two sessions in a module I will have to do the whole module again.
- I cannot do more than a total of three catch-ups over the course of the programme.
- If my progress or my behaviour in the group causes concern, I may be asked to repeat some modules of the programme or to restart it.

2. Behaviour

- I must participate in the programme to the best of my ability.
- I must not attend having consumed alcohol or illegal drugs
- I must not use language or behave in a way that is violent, racially or sexually abusive or offensive to others.
- If I do not keep to any of these requirements I may be removed from the group and could be in breach of my order or licence.

3. Evaluation

- Every session will be recorded on video to assess staff performance. Videos will focus on staff, not on group members.

- This information will help to assess the programme itself and may be used for training purposes.
- I have to fill in some questionnaires before and after the programme (and six months later). This is used to measure the effectiveness of the group and measure changes in my thoughts feelings attitudes and behaviour.

I also agree that I will:

- Maintain my goal of ending all violence and abusive behaviour towards women.
- Report to the group any violent actions, threats or impulses towards my partner or others.
- Tell staff as soon as possible if I begin a new relationship or restart a relationship with an ex-partner.
- Tell any new partner that I am on the programme and why I am on it.
- Support other group members in changing their behaviour.
- Take an active part in sessions and complete tasks set between sessions
- Not tell anyone outside the group anything about group members.

Staff has explained this form to me. I understand the programme and have agreed to take part.

Signed……

Date……

Witness……

Do not sign this form unless you understand it.

If you refuse to sign it, the Probation Officer will explain what

Will happen next.

Appendix Seven: Table displaying information gained on participants

Variable
a) Age at start of CDVP
b) Risk of Harm (medium, high or very high)
c) Ethnicity (White, Asian, Black, Mixed or Other (including Chinese)
d) Employment (yes or no)
e) Sentence type (community order, suspended sentence or custodial licence)
f) Completed CDVP (yes or no)
g) Number of months between CDVP Completion date and Order End Date
h) Number of months between CDVP suspension and Order End Date
i) If not completed CDVP; Reason for non-completion (poor attendance, poor behaviour, further offences, subjected to custody, health, employment).
j) Current IPV offence type (property damage/offensive behaviour including criminal damage, threats, disorder, affray, verbal altercation; violent offence, including common assault, ABH, GBH; harassment including breach of restraining order; sexual offence; other, including possession of a weapon, driving offences in domestic context)
k) Offence category (contact, non-contact or both)
l) Relationship to victim (ex-partner, current partner, ex-wife, wife or other including associate or other family member)
m) History of IPV (yes or no)
n) Type of IPV history record (conviction, caution or both or self-report)
o) Previous IPV category (contact, non-contact or both)
p) Number of previous IPV convictions
q) Number of previous IPV cautions
r) Previous violent convictions (yes or no)

Variable
s) Does participant have children? (yes, no or step-children)
t) Present at time of offence (yes or no)
u) Drug/alcohol at time of offence (none, alcohol, drugs or both)
v) Previous mental health issues (yes or no)
w) Marital status at time of the offence (separated/divorced, in relationship, co-habiting or married)
x) Breach of Order (yes or no)
y) If so, Reason for Breach (poor attendance, poor behaviour, failure to notify, failure to comply)

Appendix Eight: Table displaying participant recidivism variables

Variable
a) Recidivism (no, reoffended with no conviction, reconviction or both)
b) Further general non-IPV-related reconviction (yes or no)
c) Offence Type (property damage and offensive behaviour; violent offence; drug offence; burglary; other, including driving and breach).
d) Date of reconviction
e) Any further IPV conviction (yes or no)
f) IPV Offence type (classified as variable J in above table)
g) Date of first IPV offence
h) Victim type (same victim, new victim including partner ex-partner, other including new victim, associate or family member)
i) Further IPV re-offending (IPV incident not resulting in conviction) (yes or no)
j) IPV re-offence type (classified as variable k in above table)
k) Date of first IPV incident
l) Victim type (same victim, new victim including partner or ex-partner, other including new victim, associate or family member).
m) DAU Information available (yes or no)
n) Source of information for IPV re-offenders (DAU or police or self-report).

Appendix Nine: Participant Data Extraction Record for Research Project

General Participant Information

Age at start of programme:

Risk of Harm:		Low	Medium	High	Very High	
Ethnicity:						
Employed:		Yes	No			
Completed CDVP:		Yes	No			
If no, reasons why:						

Months from CDVP completion or dropout to order end date:

Offence information and history

Current Offence:			
Offence Category:	Contact	Non-Contact	Both
Type of Sentence:			
Previous Violence:	Yes	No	
IPV History:	Yes	No	
Type of IPV History	Caution	Conviction	Both

Number of previous IPV Conviction / Cautions

Breach of Order: Yes No

If yes, reasons:

Current Offence Information

Relationship to victim:

Drugs/ Alcohol at time of offence:

Mental health issues past/present: Yes No

Marital Status at time:

Have children? Yes No Step-children

Present at offence: Yes No

Recidivism Information

Did participant re-offend at any point during Order? Yes No

Type of recidivism? No conviction Conviction Both

General reconviction? (not IPV-related) Yes No

General Offence type (i.e. criminal damage):

Date of reconviction:

IPV Re-conviction? Yes No

IPV Offence Type:

Date:

Relationship to victim (i.e. same victim, new partner, other):

Further IPV re-offending? (i.e. self-reported/ arrested but did not receive a conviction):

Yes No

IPV incident type:

Note: Reported verbal abuse/ physical abuse towards partner, followed by self-report to OM is considered re-offence. Describing feelings of wanting to be abusive, but not carrying out the behaviour is not considered a re-offence.

Date of first incident:

Victim type:

DAU Information available? Yes No

Source of info for IPV reoffenders: DAU Self-report Police report

Any additional comments/notes: (for example, more than one re-offence in time period of their Order – only use information from the first; convicted and re-offended?)

_____

_____

_____

_____

_____

Appendix Ten: Table showing pre- and post- treatment means and standards deviations normed on UK IPV perpetrators (NOMS Test Battery Guide, 2012).

Psychometric Scale	Pre-Prog Mean	SD	Post-Prog Mean	SD	N
Threat to exclusive companionship	29.69	8.03	32.23	7.33	1829
Self-deprecation/envy	33.55	6.84	35.59	6.12	1830
Dependency	16.31	5.98	18.55	5.29	1870
Sexual possessiveness	7.69	3.64	8.20	3.64	1832
Competition and vindictiveness	13.76	3.11	14.84	2.93	1871
Distrust	13.24	3.97	14.35	3.58	1860

Appendix Eleven: K-S and Levene's test results for data analysis of treatment completers and non-completers.

Table i

Results of the Kolmogrov-Smirnov (K-S) tests exploring normal distribution among the 'treatment completer' and 'non-completer' groups.

Variable	Treatment Stage	Treatment Group	*D* statistic[a]	*df*	Sig.(*p*)
Age		Completer	0.098	216	.001***
		Non-completer	0.148	43	.019*
S-Ang Feeling	Pre-score	Completer	.294	216	.001***
		Non-completer	.296	43	.001***
	Post-score	Completer	.431	216	.001***
	6 month post-score	Completer	.434	50	.001***
S-Ang Verbal	Pre-score	Completer	.470	216	.001***
		Non-completer	.404	43	.001***
	Post-score	Completer	.515	216	.001***
	6 month post-score	Completer	.518	50	.001***
S-Ang Physical	Pre-score	Completer	.516	216	.001***
		Non-completer	.472	43	.001***
	Post-score	Completer	.540	216	.001***
	6 month post-score	Completer	.536	50	.001***
Total STATE	Pre-score	Completer	.257	216	.001***
		Non-completer	.259	43	.001***
	Post-score	Completer	.424	216	.001***
	6 month post-score	Completer	.447	50	.001***
T-Ang	Pre-score	Completer	.178	216	.001***
Temperament		Non-completer	.173	43	.002**
	Post-score	Completer	.155	216	.001***
	6 month post-score	Completer	.181	50	.001***
T-Ang Reaction	Pre-score	Completer	.146	216	.001***
		Non-completer	.205	43	.001***
	Post-score	Completer	.173	216	.001***

Variable	Treatment Stage	Treatment Group	*D* statistic[a]	*df*	Sig.(*p*)
	6 month post-score	Completer	.233	50	.001***
Total TRAIT	Pre-score	Completer	.103	216	.001***
		Non-completer	.158	43	.009**
	Post-score	Completer	.095	216	.001***
	6 month post-score	Completer	.159	50	.003**
AX-O	Pre-score	Completer	.184	216	.001***
		Non-completer	.159	43	.008**
	Post-score	Completer	.129	216	.001***
	6 month post-score	Completer	.160	50	.003**
AX-I	Pre-score	Completer	.135	216	.001***
		Non-completer	.116	43	.165
	Post-score	Completer	.091	216	.001***
	6 month post-score	Completer	.159	50	.003**
AC-O	Pre-score	Completer	.224	216	.001***
		Non-completer	.229	43	.001***
	Post-score	Completer	.133	216	.001***
	6 month post-score	Completer	.120	50	.070
AC-I	Pre-score	Completer	.112	216	.001***
		Non-completer	.152	43	.014*
	Post-score	Completer	.074	216	.001***
	6 month post-score	Completer	.151	50	.006**
Anger Expression	Pre-score	Completer	.178	216	.001***
Index		Non-completer	.187	43	.001***
	Post-score	Completer	.116	216	.001***
	6 month post-score	Completer	.185	50	.001***
ECR-R Anxious	Pre-score	Completer	.084	216	.001***
		Non-completer	.087	43	.200
	Post-score	Completer	.056	216	.091
	6 month post-score	Completer	.143	50	.012*
ECR-R Avoidant	Pre-score	Completer	.053	216	.200
		Non-completer	.060	43	.200
	Post-score	Completer	.059	216	.068
	6 month post-score	Completer	.110	50	.176
RATOS	Pre-score	Completer	.066	213[b]	.023*

Variable	Treatment Stage	Treatment Group	*D* statistic[a]	*df*	Sig.(*p*)
		Non-completer	.086	43	.200
	Post-score	Completer	.048	213	.200
	6 month post-score	Completer	.089	50	.200
IRS Threat	Pre-score	Completer	.095	216	.001***
		Non-completer	.130	43	.065
	Post-score	Completer	.081	216	.002**
	6 month post-score	Completer	.159	49[c]	.003**
IRS Envy	Pre-score	Completer	.098	216	.001***
		Non-completer	.163	43	.006**
	Post-score	Completer	.130	216	.001***
	6 month post-score	Completer	.198	49	.001***
IRS Dependent	Pre-score	Completer	.105	216	.001***
		Non-completer	.187	43	.001***
	Post-score	Completer	.135	216	.001***
	6 month post-score	Completer	.185	49	.001***
IRS Sexual	Pre-score	Completer	.121	216	.001***
Possessiveness		Non-completer	.151	43	.015*
	Post-score	Completer	.188	216	.001***
	6 month post-score	Completer	.212	49	.001***
IRS Competitive	Pre-score	Completer	.113	216	.001***
		Non-completer	.140	43	.033*
	Post-score	Completer	.126	216	.001***
	6 month post-score	Completer	.162	49	.002**
IRS Distrust	Pre-score	Completer	.136	216	.001***
		Non-completer	.150	43	.016*
	Post-score	Completer	.119	216	.001***
	6 month post-score	Completer	.186	49	.001***

[a]The test statistic for the K-S test is denoted by D. [b]There were missing data on three different participants on the RATOS. [c]There were missing data for one participant on the follow up IRS. * $p < .05$. ** $p < .01$. *** $p \leq .001$.

Table ii

Table showing results of the K-S Test on date variables.

Variable	Treatment Group	D statistic[a]	df	Sig.(p)
Number of months from Group End to Order Expiry Date	Completer	.098	216	.012*
No of months from Suspend date to Order Expiry Date	Non-completer	.294	43	.000***
No. of months to 'Overall IPV Recidivism'	Completer	.116	45	.156
	Non-completer	.175	11	.200
No. of months to first non-IPV reconviction	Completer	.270	18	.001***
	Non-completer	.146	9	.200
No. of months to first IPV reconviction	Completer	.155	12	.200
	Non-completer	.151	9	.200
No. of months to first IPV alleged reoffence	Completer	.122	35	.200
	Non-completer	.343	6	.026*

Table iii

Table showing Homogeneity of variance between treatment completers and non-completers groups.

Variable	*F*	*df*	*p*
No. of mths to Overall IPV Recidivism	6.22	1, 54	.016**
No. of mths to non-IPV reconviction	1.03	1, 25	.319
No. of mths to first IPV reconviction	2.28	1, 19	.148
No. of months to first IPV alleged reoffending	.051	1, 39	.822
Age	1.65	1, 257	.200
S-Ang Feeling	1.71*	1, 257	.193
S-Ang Verbal	3.54	1, 257	.061
S-Ang Physical	2.73	1, 257	.100
Total STATE	2.48	1, 257	.117
T-Ang Temperament	.91	1, 257	.341

T-Ang Reaction	.62	1, 257	.430
Total TRAIT	.59	1, 257	.442
AX-O	.39	1, 257	.535
AX-I	.59	1, 257	.442
AC-O	.39	1, 257	.534
AC-I	.01	1, 257	.928
Anger Exp Index	.11	1, 257	.738
ECR-R Anxious	1.33	1, 257	.251
ECR-R Avoidant	.04	1, 257	.838
RATOS	.28	1, 254	.597
IRS Threat	1.55*	1, 257	.214
IRS Envy	1.86*	1, 257	.173
IRS Dependent	.40*	1, 257	.525
IRS Sexual Possessiveness	3.48*	1, 257	.063
IRS Competitive	6.36*	1, 257	.012**
IRS Distrust	1.74*	1, 257	.189

* Levene's Based on Median ** significant at $p < .05$.

Appendix Twelve: Table showing results from Friedman's ANOVA

Variable	Treatment stage	*M (SD)*	*Mdn*	*M Rank*	*Df*	χ^2	*P*	*W*
S-Ang/F	Pre	51.88 (16.93)	40.00	2.31	2	17.11	.001**	.171
	Post	44.00 (13.40)	40.00	1.87				
	Follow-up	42.40 (8.28)	40.00	1.82				
S-Ang/V	Pre	58.46 (15.15)	50.00	2.18	2	9.57	.008	.096
	Post	54.14 (11.95)	50.00	1.92				
	Follow-up	53.40 (9.66)	50.00	1.90				
S-Ang/P	Pre	55.34 (13.84)	50.00	2.12	2	8.96	.009	.090
	Post	52.62 (10.59)	50.00	1.96				
	Follow-up	50.90 (6.36)	50.00	1.92				
S-Ang	Pre	52.72 (17.84)	42.50	2.34	2	17.96	.001**	.180
	Post	44.56 (14.16)	40.00	1.85				
	Follow-up	43.70 (11.47)	40.00	1.81				
T-Ang/T	Pre	68.56 (24.68)	75.00	2.34	2	14.22	.001*	.142
	Post	59.92 (25.00)	70.00	1.94				
	Follow-up	54.26 (24.27)	52.50	1.72				
T-Ang/R	Pre	37.72 (28.62)	30.00	2.30	2	10.69	.004	.107
	Post	30.86 (25.55)	25.00	1.95				
	Follow-up	28.14 (28.22)	20.00	1.75				
T-Ang	Pre	55.48 (30.28)	60.00	2.42	2	18.30	.001**	.183
	Post	44.30 (29.37)	40.00	1.93				
	Follow-up	37.92 (29.81)	32.50	1.65				
AX-O	Pre	60.38 (32.28)	70.00	2.37	2	14.60	.001*	.146
	Post	48.66 (31.77)	55.00	1.98				
	Follow-up	39.80 (32.04)	30.00	1.65				
AX-I	Pre	52.88 (31.46)	55.00	2.26	2	7.02	.032	.070

	Post	45.98 (28.30)	45.00	1.98				
	Follow-up	39.38 (30.21)	35.00	1.76				
AC-O[a]	Pre	34.52 (28.97)	25.00	1.51	2	19.97	.001**	.200
	Post	52.26 (29.15)	50.00	2.27				
	Follow-up	56.94 (27.96)	60.00	2.22				
AC-I[a]	Pre	40.06 (22.95)	40.00	1.48	2	23.75	.001**	.238
	Post	57.90 (22.61)	60.00	2.13				
	Follow-up	66.32 (24.58)	72.50	2.39				
AX Index	Pre	61.98 (30.33)	72.50	2.49	2	23.42	.001**	.234
	Post	43.06 (29.92)	42.50	1.92				
	Follow-up	32.80 (29.75)	25.00	1.59				
IRS Threat[b]	Pre	30.00 (7.52)	29.00	1.60	2	21.32	.001**	.217
	Post	32.84 (7.21)	34.00	1.92				
	Follow-up	34.47 (7.55)	37.00	2.48				
IRS Envy[b]	Pre	32.98 (6.12)	33.00	1.78	2	4.65	.100	.047
	Post	34.94 (5.99)	36.00	2.06				
	Follow-up	35.39 (6.93)	38.00	2.16				
IRS Dependent[b]	Pre	16.78 (6.25)	18.00	1.73	2	13.29	.001*	.136
	Post	18.20 (5.66)	19.00	1.89				
	Follow-up	19.45 (5.09)	21.00	2.38				
IRS Sexual Possessive[b]	Pre	8.08 (3.66)	8.00	1.78	2	4.32	.115	.044
	Post	8.86 (3.93)	8.00	2.11				
	Follow-up	8.98 (3.70)	8.00	2.11				
IRS Competitive[b]	Pre	13.88 (3.07)	15.00	1.63	2	13.66	.001*	.139
	Post	14.92 (2.69)	15.00	2.07				
	Follow-up	15.22 (2.82)	16.00	2.30				
IRS Distrust[b]	Pre	12.51 (4.32)	13.00	1.68	2	8.94	.011	.091

	Post	14.24 (3.10)	15.00	2.19				
	Follow-up	13.69 (3.96)	15.00	2.12				
ECR-R Anxious[c]	Pre	3.43 (1.06)	3.59	2.27	2	7.07	.029	.071
	Post	3.05 (.93)	2.89	1.97				
	Follow-up	2.92 (.97)	2.67	1.76				

Note. N = 50 on STAXI-II and ECR-R at follow-up stage. N = 49 on IRS at follow-up stage. [a] Higher scores on AC-O and AC-I indicate higher levels of reported Anger Control. [b] Higher scores on the IRS indicate lower levels of jealousy in that domain. [c]Lower scores on the ECR-R indicate more secure attachment. *significant at p = .001. **significant at p < .001.

Appendix Thirteen: Table showing recidivists versus non-recidivists frequencies and Chi-Square results

Variable	Treatment Completers				N	Non-completers				N	χ^{2a}
IPV and Non-IPV offending combined	No 161 (74.1)	Re-offend 31 (14.4)	Re-convict 20 (9.3)	Both 5 (2.3)	216	No 26 (60.5)	Re-offend 2 (4.7)	Re-convict 11 (25.6)	Both 4 (9.3)	43	16.59***
Non-IPV reconviction	Yes 17 (7.9)	No 199 (92.1)			216	Yes 9 (20.9)	No 34 (79.1)			43	FET*
IPV reconviction	Yes 12 (5.6)	No 204 (94.4)			216	Yes 9 (20.9)	No 34 (79.1)			43	FET***
IPV reconviction victim	Same victim 11 (91.7)	New victim 0 (0)	Other 1 (8.3)		12	Same victim 7 (77.8)	New victim 2 (22.2)	Other 0 (0)		9	FET 3.05
IPV alleged reoffending (no conviction)	Yes 35 (16.2)	No 181 (83.8)			216	Yes 6 (14.0)	No 37 (86.0)			43	.136
IPV alleged reoffending victim	Same victim 28 (80.0)	New victim 5 (14.3)	Other 2 (5.7)		35	Same victim 5 (83.3)	New victim 1 (16.7)	Other 0 (0)		6	FET .516
DAU info available	Yes 91 (42.1)	No 125 (57.9)			216	Yes 22 (51.2)	No 21 (48.8)			43	1.19
Source of info for re-offend	DAU/Police 16 (45.7)	Self-report 19 (54.3)			35	DAU/Police 3 (50)	Self-report 3 (50)			6	FET

Appendix Fourteen: K-S and Levene's test results for data analysis of recidivists and non-recidivists.

Table i

Table showing results of Kolmogrov-Smirnov (K-S) tests exploring normal distribution among the 'recidivist' and 'non-recidivist' groups (within treatment completers at the post-treatment stage).

Variable	Group	*M (SD)*	*Mdn*	*D*	*df*	*P*
Age	Recidivist	33.42 (8.58)	31.00	.144	45	.020*
	Non-recidivist	36.75 (11.12)	36.00	.092	171	.001***
S-Ang/F	Recidivist	41.64 (13.09)	40.00	.439	45	.001***
	Non-recidivist	42.24 (11.78)	40.00	.429	171	.001***
S-Ang/V	Recidivist	53.98 (11.72)	50.00	.552	45	.001***
	Non-recidivist	54.16 (11.49)	50.00	.513	171	.001***
S-Ang/P	Recidivist	52.87 (10.90)	50.00	.537	45	.001***
	Non-recidivist	51.95 (8.93)	50.00	.540	171	.001***
S-Ang	Recidivist	43.09 (14.62)	40.00	.428	45	.001***
	Non-recidivist	43.37 (13.17)	40.00	.423	171	.001***
T-Ang/T	Recidivist	62.80 (23.30)	70.00	.141	45	.026*
	Non-recidivist	63.00 (25.29)	70.00	.159	171	.001***
T/Ang/R	Recidivist	32.58 (25.11)	30.00	.149	45	.014*
	Non-recidivist	31.89 (26.47)	20.00	.188	171	.001***
T-Ang	Recidivist	49.36 (27.96)	50.00	.097	45	.200
	Non-recidivist	49.06 (28.96)	50.00	.098	171	.001***
AX-O	Recidivist	53.02 (29.23)	45.00	.130	45	.056
	Non-recidivist	53.09 (31.85)	55.00	.128	171	.001***
AX-I	Recidivist	48.76 (30.76)	50.00	.125	45	.075
	Non-recidivist	50.04 (28.15)	50.00	.097	171	.001***
AC-O	Recidivist	46.24 (30.00)	40.00	.161	45	.005**
	Non-recidivist	43.00 (28.99)	40.00	.124	171	.001***
AC-I	Recidivist	60.89 (21.14)	60.00	.103	45	.200
	Non-recidivist	54.51 (24.79)	55.00	.075	171	.020*
AX Index	Recidivist	45.76 (29.87)	45.00	.182	45	.001***
	Non-recidivist	51.11 (29.47)	50.00	.108	171	.001***

ECR Anxious	Recidivist	2.95 (1.08)	2.83	.064	45	.200
	Non-recidivist	3.02 (.990)	2.94	.070	171	.038*
ECR Avoidant	Recidivist	2.68 (.737)	2.61	.084	45	.200
	Non-recidivist	2.92 (.890)	2.89	.061	171	.200
RATOS	Recidivist	32.00 (11.96)	33.00	.094	44	.200
	Non-recidivist	34.14 (11.72)	35.00	.047	169	.200
IRS Threat	Recidivist	30.62 (3.69)	31.00	.101	45	.200
	Non-recidivist	32.20 (7.41)	33.00	.093	171	.001***
IRS Envy	Recidivist	35.47 (5.62)	37.00	.175	45	.001***
	Non-recidivist	34.09 (7.04)	35.00	.131	171	.001***
IRS Dependent	Recidivist	17.93 (4.73)	18.00	.198	45	.031*
	Non-recidivist	18.33 (5.36)	20.00	.145	171	.001***
IRS Sexual Possessive	Recidivist	7.53 (3.60)	7.00	.182	45	.001***
	Non-recidivist	8.16 (3.50)	8.00	.191	171	.001***
IRS Competitive	Recidivist	14.16 (2.91)	14.00	.137	45	.033*
	Non-recidivist	14.26 (3.09)	15.00	.132	171	.001***
IRS Distrust	Recidivist	13.71 (3.15)	13.00	.145	45	.019*
	Non-recidivist	14.13 (3.16)	15.00	.140	171	.001***

* significant at < .05. ** significant at < .01 *** significant at ≤ .001.

Table ii

Table showing homogeneity of variance on post-treatment psychometric scores and Age amongst 'recidivists' and 'non-recidivists'.

Variable	***F***	***Df (1)***	***Df (2)***	***p***
Age	5.40	1	214	.021*
S-Ang/F	.012	1	214	.914
S-Ang/V	.009	1	214	.923
S-Ang/P	.339	1	214	.561
S-Ang	.041	1	214	.839
T-Ang/T	.713	1	214	.399
T/Ang/R	.089	1	214	.765
T-Ang	.396	1	214	.530
AX-O	1.41	1	214	.236
AX-I	.468	1	214	.495
AC-O	.316	1	214	.575
AC-I	1.18	1	214	.279
AX Index	.005	1	214	.943
ECR Anxious	.190	1	214	.663
ECR Avoidant[a]	2.40	1	214	.213
RATOS	.281	1	211	.597
IRS Threat	.651	1	214	.421
IRS Envy	4.83	1	214	.029*
IRS Dependent	.615	1	214	.434
IRS Sexual Possessive	.062	1	214	.804
IRS Competitive	.075	1	214	.785
IRS Distrust	.003	1	214	.957

Note. Levene's based on Median unless stated by following K-S test finding skewness of data in at least one of the groups. [a] Levene's based on Mean. * significant at < .05

Appendix Fifteen: Table showing Mann Whitney U outputs of treatment completer recidivists and non-recidivists.

Variable	Group	*N*	*M (SD)*	*Mdn*	*U*	*z*	*p*	*r*
Age	Recidivist	45	33.42 (8.58)	31.00	3258.50	-1.58	.111	-.107*
	Non-recidivist	171	36.75 (11.12)	36.00				
S-Ang/F	Recidivist	45	41.64 (13.09)	40.00	3500.00	-1.04	.303	-.071
	Non-recidivist	171	42.24 (11.78)	40.00				
S-Ang/V	Recidivist	45	53.98 (11.72)	50.00	3794.50	-.247	.823	-.017
	Non-recidivist	171	54.16 (11.49)	50.00				
S-Ang/P	Recidivist	45	52.87 (10.90)	50.00	3768.00	-.559	.544	-.038
	Non-recidivist	171	51.95 (8.93)	50.00				
S-Ang	Recidivist	45	43.09 (14.62)	40.00	3533.50	-.927	.350	-.063
	Non-recidivist	171	43.37 (13.17)	40.00				
T-Ang/T	Recidivist	45	62.80 (23.30)	70.00	3801.00	-.126	.901	-.009
	Non-recidivist	171	63.00 (25.29)	70.00				
T-Ang/R	Recidivist	45	32.58 (25.11)	30.00	3713.50	-.362	.718	-.02
	Non-recidivist	171	31.89 (26.47)	20.00				
T-Ang	Recidivist	45	49.36 (27.96)	50.00	3830.00	-.047	.962	-.003
	Non-recidivist	171	49.06 (28.96)	50.00				

AX-O	Recidivist	45	53.02 (29.23)	45.00	3839.00	-.022	.984	-.001
	Non-recidivist	171	53.09 (31.85)	55.00				
AX-I	Recidivist	45	48.76 (30.76)	50.00	3707.00	-.378	.706	-.026
	Non-recidivist	171	50.04 (28.15)	50.00				
AC-O	Recidivist	45	46.24 (30.00)	40.00	3639.00	-.561	.583	-.038
	Non-recidivist	171	43.00 (28.99)	40.00				
AC-I	Recidivist	45	60.89 (21.14)	60.00	3313.00	-1.44	.153	-.098
	Non-recidivist	171	54.51 (24.79)	55.00				
AX Index	Recidivist	45	45.76 (29.87)	45.00	3476.50	-.996	.320	-.067
	Non-recidivist	171	51.11 (29.47)	50.00				
ECR Anxious	Recidivist	45	2.95 (1.08)	2.83	3706.50	-.378	.702	-.026
	Non-recidivist	171	3.02 (.990)	2.94				
IRS Threat	Recidivist	45	30.62 (3.69)	31.00	3210.00	-1.71	.086	-.116*
	Non-recidivist	171	32.20 (7.41)	33.00				
IRS Envy	Recidivist	45	35.47 (5.62)	37.00	3483.00	-.981	.332	-.067
	Non-recidivist	171	34.09 (7.04)	35.00				
IRS Dependent	Recidivist	45	17.93 (4.73)	18.00	3512.00	-.904	.362	-.061
	Non-recidivist	171	18.33 (5.36)	20.00				

IRS Sexual Possessive	Recidivist	45	7.53 (3.60)	7.00	3369.00	-1.30	.188	-.088
	Non-recidivist	171	8.16 (3.50)	8.00				
IRS Competitive	Recidivist	45	14.16 (2.91)	14.00	3713.00	-.364	.712	-.025
	Non-recidivist	171	14.26 (3.09)	15.00				
IRS Distrust	Recidivist	45	13.71 (3.15)	13.00	3498.50	-.942	.346	-.341
	Non-recidivist	171	14.13 (3.16)	15.00				

* small effect size.

Appendix Sixteen: A table showing differences between 'treatment completer' recidivists and non-recidivists based on Overall IPV recidivism.

Variable		**Overall IPV Recidivism –No (N=171)**					**Overall IPV Recidivism - Yes (N=45)**					χ^2	*p*
N (%)													
Risk of Harm	Medium	High					Medium	High				FET	.175
	162 (94.7)	9 (5.3)					40 (88.9)	5 (11.1)					
Ethnicity[a]	White	Asian	Black	Mixed	Other		White	Asian	Black	Mixed	Other	FET	.518
	134 (82.7)	7 (4.3)	14 (8.6)	6 (3.7)	1 (1.6)		35 (83.3)	1 (2.4)	2 (4.8)	3 (7.1)	1 (2.4)	3.21	
Employment[b]	Yes	No					Yes	No				.010	.921
	99 (60.4)	65 (39.6)					25 (59.5)	17 (40.5)					
Sentence type	Community	Susp Sent	Licence				Comm.	Susp Sent	Licence			.983	.612
	120 (70.2)	45 (26.3)	6 (3.5)				29 (62.8)	15 (33.3)	1 (2.2)				
Offence	Contact	Non-contact	Dual				Contact	Non-contact	Dual			.1.62	.444
category	104 (60.8)	26 (15.2)	41 (24.0)				24 (53.3)	6 (13.3)	15 (33.3)				
Relationship	Ex-partner	Current	Wife	Ex-wife	Other		Ex-partner	Current	Wife	Ex-wife	Other	5.46	.243
to victim		Partner						partner					
	44 (25.7)	78 (45.6)	34 (19.9)	7 (4.1)	8 (4.7)		8 (17.8)	28 (62.2)	8 (17.8)	1 (2.2)	0 (0)		
IPV history	Yes	No					Yes	No				2.82	.093
	148 (86.5)	23 (13.5)					43 (95.6)	2 (14.4)					
IPV Record[c]	Convictions	Cautions	Both	Self-report			Convictions	Cautions	Both	Self-report		3.80	.284
		38 (25.7)	8 (5.4)	62 (41.9)				12 (28.6)	5 (11.9)				

	40 (27.0)				13 (31.0)			12 (28.6)		
IPV history offence[d] **category**	Contact 56 (38.1)	Non-contact 29 (19.7)	Dual 62 (42.2)		Contact 11 (26.2)	Non-contact 9 (21.4)	Dual 22 (52.4)		2.19	.347
Previous violence	Yes: 58 (33.9)	No: 113 (66.1)			Yes 20 (44.4)	No 25 (55.6)			.191	.191
With children	Yes: 103 (30.2)	No: 49 (28.7)	Step-children 19 (11.1)		Yes 30 (66.7)	No 7 (15.6)	Step-children 8 (17.8)		3.86	.145
Child Present at offence[e]	Yes 80 (65.6)	No 42 (34.4)			Yes 27 (71.1)	No 11 (28.9)			.393	.531
Substance misuse	No 46 (26.9)	Alcohol 103 (60.2)	Drugs 9 (5.3)	Both 13 (7.6)	No 13 (28.9)	Alcohol 25 (55.6)	Drugs 2 (4.4)	Both 5 (11.1)	FET .948	.837
History of mental health	Yes 82 (48.0)	No 89 (52.0)			Yes 19 (42.2)	No 26 (57.8)			.470	.493
Marital status	Separated 54 (31.6)	In rel 38 (22.2)	Co-habiting 46 (26.9)	Married 33 (19.3)	Separated 10 (22.2)	In rel 20 (44.4)	Co-habiting 6 (13.3)	Married 9 (20.0)	10.34	.016*
Breach	Yes 5 (2.9)	No 166 (97.1)			Yes 4 (8.9)	No 41 (91.1)			FET	.093

Note. [a] IPV Overall Re-offending = No, N = 162; IPV Overall Re-offending = Yes, N = 42 due to missing data. [b] IPV Overall Re-offending No, N = 164, IPV Overall Re-offending, N = 42 due to missing data. [c] IPV Overall Re-offending = No, N = 148; IPV Overall Re-offending = Yes, N = 42 due to data being non-applicable to those with no IPV history. [d] IPV Overall Re-offending = No, N = 147, IPV Overall Re-offending = Yes, N = 42 due to data being non-applicable to those with no IPV history or missing data. [e] IPV Overall Re-offending = No, N = 122; IPV Overall Re-offending = Yes, N = 38 due to non-applicable or missing data. * significant at $p < .05$.