

Changing for Whom?

The Role of Existential Motivation, Couple Participation and Relationship Context Disagreement in Men's Domestic Violence Perpetrator Programmes

Name and initials: Downes R.E.
Student number: S2556405
E-mail address: r.e.downes@student.rug.nl
Cohort: 2018-2020
First assessor: prof. dr. R. Veenstra
Second assessor: prof. dr. A. Lichtwarck-Aschoff
Programme: Research Master Behavioural and Social Sciences
Track: Lifespan Development and Socialisation
ECs: 30
Date: 07-02-2022
Word count: 9062

Are there deviations of the Master's thesis from the proposed plan?

No

Yes, please explain below the deviations

Initially I planned to examine three outcome variables: emotional violence, physical and sexual violence. Because of shortage of space and because emotional violence is reported most, I only tested one of those, emotional violence, through multilevel analysis.

Abstract

In this study, motivation and couples' relationship context disagreement (about relationship status and relationship hopes) were used to predict emotional violence. Secondary analyses were done on longitudinal data among male participants from European programmes. Perpetrators' self-reports were compared with (ex-)partner reports about the perpetrators' behaviour. To assess the value of (ex-)partner participation and relationship context disagreement, treatment impact was compared for perpetrators who participated solo versus perpetrators who participated with an (ex-)partner and (dis)agreed. Finally, patterns of attrition were examined.

Results revealed that emotional violence significantly decreased over time. Clients with higher existential motivation had higher baseline scores in emotional violence, but their use of violence also decreased significantly more over time. Clients whose partner participated and who disagreed with their partner about the relationship context had higher average scores on emotional violence than solo participating clients. Results did not support lower treatment impact for clients who disagreed with their partner about the relationship context.

Challenges in analysis resulting from attrition and variation in administration, provide directions for improvement and growth. The role of existential motivation and relationship context disagreement deserve further investigation and are suitable for multi-site, multi-country evaluations of perpetrator programmes through multilevel analysis.

Table of contents

Abstract.....	3
List of figures.....	5
List of tables.....	5
1. Introduction.....	6
1.1 Intimate partner violence: prevalence, consequences and services	6
1.2 The current study	7
2. Theory.....	13
2.1 Motivation.....	13
Existential versus functional motivation.....	13
Motivation is no panacea	14
2.2 Relationship context.....	15
The importance of couple reports	15
Mechanisms in dyads: from expectations, preferences and disagreement to IPV	15
2.3 Research on programme participation, relationship context and motivation	16
2.4 Implicit implications of programme participation	18
3. Method.....	19
3.1 Procedure	19
3.1 Participants.....	20
3.2 Instrument	24
3.3 Measures	25
3.5 Data analysis procedure	28
Descriptive statistics	28
Multilevel modelling.....	28
4. Results.....	30
4.1 Descriptive statistics	30
4.2 Multilevel analysis	33
Subject- and time-level effects on emotional violence	33
Summary findings.....	44
5. Discussion and conclusion.....	45
Reference list	50
Appendix 1 - Intimate partner violence estimates	61
1. Intimate partner violence victimisation rates	61
2. Service provision to victims and perpetrators.....	63
2.1 Residential care for victims	63
2.2 Non-residential care for victims.....	64
2.3 Perpetrator programmes.....	65
2.4 European context of perpetrators programmes	65
2.5 Availability and obstacles of programmes in countries under study	66
Appendix 2 - Attrition.....	69
Attrition in general.....	69
Present study and problems with attrition.....	70
Participation rates across organisations	70
Results.....	73
Discussion.....	75
Appendix 3 - Factor analysis	77

List of figures

Figure 1 . Conceptual framework of the effects of motivation and relationship disagreement on emotional violence.....	12
Figure 2 . Normal scores against time-level standardised residuals, for emotional violence untransformed (left) and log-10 transformed (right).	34
Figure 3 . Normal scores against subject-level standardised residuals, for emotional violence untransformed (left) and log-10 transformed (right), null-model.	35
Figure 4 . Normal scores against subject-level standardised residuals of emotional	43
Figure 5 . Subject-level residuals for emotional violence (standardised residuals against fixed part prediction), full model.....	44
Figure 6 . Caterpillar plots for ranked subject-level residuals.	44

List of tables

Table 1 . Participation rates for clients across all timepoints.	20
Table 2 . Participation rates for (ex-)partners across all timepoints.	20
Table 3 . Participation of male clients per organisation, across all timepoints....	21
Table 4 . Participation of female (ex-)partners per organisation, across all timepoints.....	21
Table 5 . Demographic statistics (age, referral route, employment and income) of male clients at baseline.	23
Table 6 . Clients' relationship status and hopes, recoded	30
Table 7 . (Ex-)Partners' relationship status and hopes, recoded.....	31
Table 8 . Clients' relationship context disagreement at baseline and follow-up.	31
Table 9 . Clients' mean frequency of emotional violence according clients and (ex-)partners.....	32
Table 10 . Mean frequency of emotional violence and differences between linked partners from 15 couples.....	33
Table 11 . Results of multilevel analysis for emotional violence, fixed effects. .	37
Table 12 . Results of multilevel analysis with predictors as random effects.	41

1. Introduction

1.1 Intimate partner violence: prevalence, consequences and services

As intimate partner violence (IPV) is widespread in its prevalence and consequences, insufficient and ineffective service provision to victims and perpetrators call for improvement. Multi-site, multi-country studies such as the current study, can contribute to such improvements, by examining the effectiveness of perpetrator programmes (Scambor et al., 2014). Issues of programmes' effectiveness in reducing perpetrators' abuse are tied in with questions about motivation, study design, and attrition (Donovan & Griffiths, 2015). For the countries under study - Bulgaria, UK, Croatia and Italy - between 13 and 25 per cent of women experience IPV in their lifetime, versus 2 to 19 per cent in the past years, as estimated by international surveys (FRA - European Union Agency for Fundamental Rights, 2014; World Health Organization, 2013) and national surveys (Elkin, 2019; Istat 2015; Ivanova, 2016). See Appendix 1 for a discussion of specific estimates. Whereas the frequency and severity of abuse is high in population-based samples, it is even worse in facility-based samples (Walby, 2005).

The consequences are devastating for victims, children, bystanders, and society at large. Victims suffer both short- and long-term consequences, varying from isolation, psychological issues, negative health outcomes, trauma, physical injuries to femicide (Campbell, 2002; Corradi et al., 2018; Dolezal et al., 2009; Westmarland & Kelly, 2013). Children who witness domestic violence suffer from lower educational achievement, antisocial problems, mental health issues, psycho-somatic complaints and difficulties in intimate relationships (Jaffe et al., 1990; Kitzmann et al., 2003; Siegel, 2013). On the long term, they run a higher risk of becoming a victim or perpetrator later in life (Thornberry et al., 2012; Widom, 1989). These consequences result in large estimated health care costs and societal costs, although the extent and impact of the

abuse and the costs remain a large dark figure (Oliver et al., 2019; Walby, 2004). Not intervening in perpetration comes with even greater costs, which justifies improvement, prevention and breaking cycles of abuse (Hester & Lilley, 2014; Walby, 2004).

The implications of these figures are worrisome, as victims are not only failed because they remain undiscovered, but abuse also does not stop once support services are present. Both perpetrators and victims experience barriers in disclosing abuse and seeking help (Morgan et al., 2014). As a result, two-thirds of the victims do not reach out to professional support (FRA, 2014). Victims who *do* receive help, endure between two to three years of abuse before they seek help (SafeLives, 2015). Most victims seek professional help five times in the year before receiving effective help and undergo 50 incidents before being helped effectively (SafeLives, 2015).

Unfortunately, service provision to victims and perpetrators continues to be lacking in quantity and quality in the European Union and the four countries under study. The number of women's shelters and victim services do not meet the required minimum agreed upon by European countries through the Istanbul Convention in 2011. Perpetrator services in Bulgaria, Croatia, Italy and to a certain extent the UK, deal with issues in legislative ratification, community recognition and funding. This causes a deficit in programmes offered, professionals being trained and perpetrators that can be reached. Appendix 1 provides a background on victim and perpetrator services in the E.U.

1.2 The current study

This study focuses on the effectiveness of perpetrator programmes, which are designed and based upon the assumption that abusive behaviours can be unlearned and perpetrators should take responsibility for the abuse (Lilley-Walker et al., 2016). There

are reasons to critically assess this assumption. First, research on perpetrator programmes shows limited or insignificant effects. This can be explained by either methodological shortcomings or inherent ineffectiveness. Methodological limitations include the way data are interpreted, the lack of comparison groups, limited outcomes such as police call-outs, small sample sizes and not reporting attrition (Lilley-Walker et al., 2016; Feder & Wilson, 2005; Westmarland & Kelly, 2013). Limited or insignificant effects may also indicate programmes are truly ineffective, which can be difficult to accept “when we are in the business of funding, developing, delivering, or evaluating perpetrator programmes” (McGinn et al., 2019).

Second, although many victims feel supported and empowered because their (ex-)partner is in a programme (Chung et al. 2020), such programmes also have unintended adverse consequences for victims (McGinn et al., 2019; Westmarland & Kelly, 2013; 2020; Ginés et al. 2015). “Service design and practice might overlook, exclude or restrict the safety and support needs of certain groups of victims/survivors (more than others) and thus undermine efforts towards perpetrator accountability” (Chung et al., 2020, p.15). These two issues - limited effectiveness and potential adverse consequences - call for a re-prioritisation of victims’ safety through improvement of research and practice.

Perpetrator programmes can benefit from a better understanding of how programme participation affects victims negatively, by taking into account how perpetrators’ motivations to participate and their relationship context affect treatment impact. So far, this interdependence remains understudied (e.g. Crane et al., 2013; Downes & Vall, 2020; Gray et al., 2016). This study seeks to discover whether treatment impact (the decrease of emotional violence over time) differ among male participants of perpetrator programmes. The effects are predicted by 1) clients’

motivation to participate in the programme, 2) participation as couple and 3) disagreement with their (ex-)partner about the relationship status and their hopes for the future of the relationship. These features are captured in the main research question: *Do existential motivation, couple participation, and relationship context disagreement, increase treatment impact of IPV perpetrators?* Its hypotheses are explained below and presented in Figure 1.

Research has shown the beneficial effects of existential over functional motivation on treatment impact (McGinn et al., 2020). Existential here means that perpetrators are motivated to participate because they want to be a better parent, partner or father. Functional refers to wanting to participate in order to fulfil an external objective, such as avoiding a court sentence or winning back one's partner. This leads to the first hypothesis: *H1: Emotional violence would decrease more over time for perpetrators with higher existential motivation.*

Treatment impact may however be altered depending on ecological factors, such as relationship dynamics (Stith et al., 2012). The effects of a relationship are measured here through couple participation and couples' disagreement about the relationship context. Couple participation means both the perpetrator and (ex-)partner participate. (Ex-)Partner participation is an indicator of victim support, offered to women and children by perpetrator programmes (Chung et al., 2020). The presence of partner participation and partner contact promotes perpetrators and programmes to act in the interest and safety of the victims (Chung et al., 2020; Day et al., 2019). By examining partner participation, a distinction can be made whether treatment impact is higher for perpetrators whose partner participates, versus those whose partner does not participate. Mere couple participation is no sufficient measure, more so the content and quality of the relationship. Research has shown that IPV can be predicted

from relationship dissatisfaction (Stith et al., 2008), non-acceptance of the relationship status (Crane et al. 2013), and relationship separation (Walby & Allen, 2004). Also, participation in perpetrator programmes can be motivated by not wanting to lose one's partner (Hester et al., 2006), wanting to win back one's partner, and motivation diminishing as a result of separation (Gray et al., 2016). In such cases of disagreement, partner participation is not thought to be a more beneficial effect than solo participation in regards of treatment impact. Therefore, the role is examined of couples' disagreement about the current relationship and the hopes for the future of the relationship. This leads to the following hypotheses:

H2: Perpetrators who agree with their (ex-)partner about the relationship context, would report a stronger decrease of emotional violence over time, than perpetrators who participate solo.

H3: Perpetrators who participate solo would report a stronger decrease of emotional violence than perpetrators who disagree with their (ex-)partner about the relationship context.

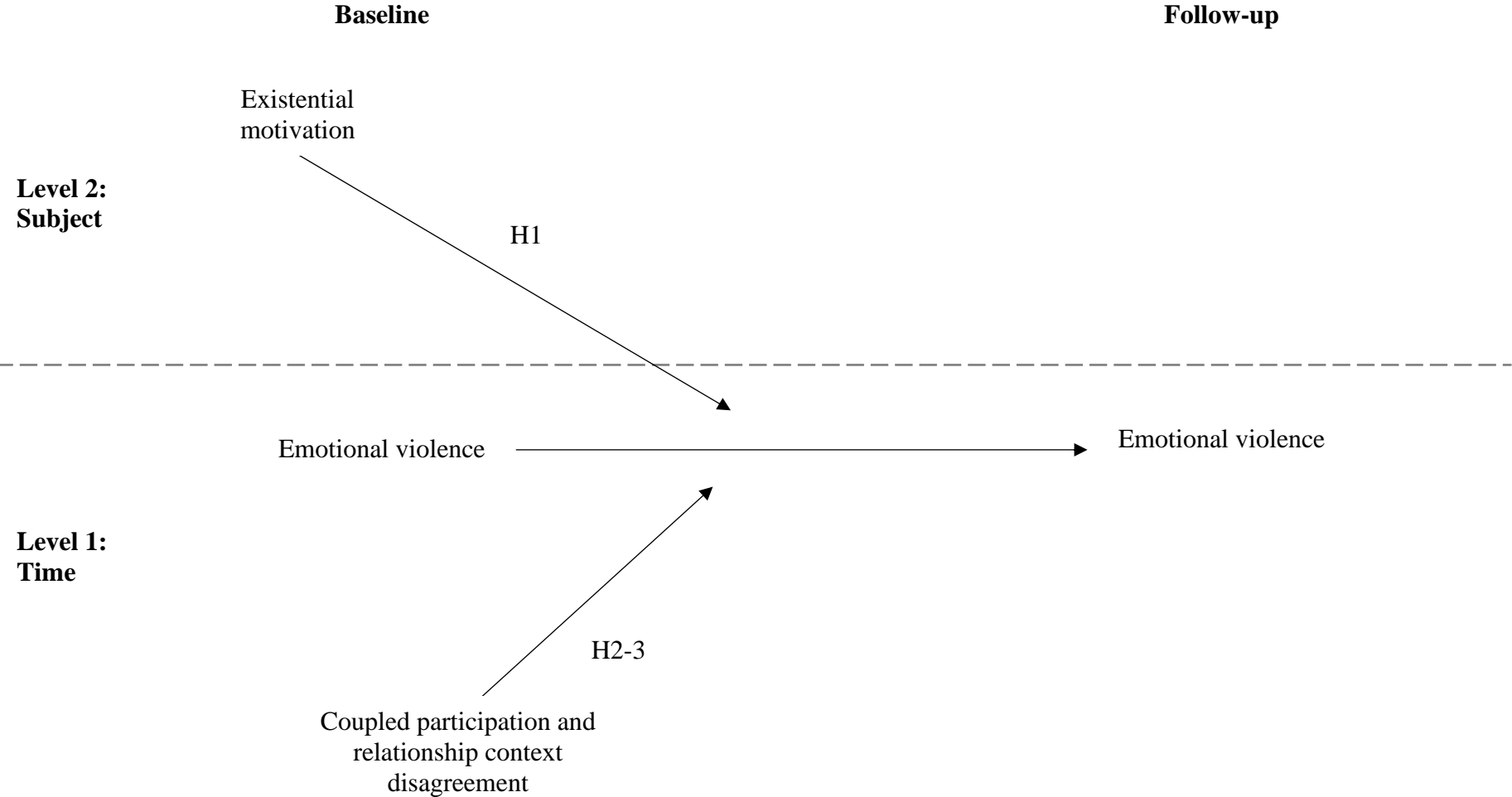
Furthermore, it matters *who* reports success of a programme: the perpetrator or the victim. As this study is guided by the role, needs, and safety of the victims in the context of perpetrator programmes, comparing victims' perspectives to those of perpetrators is essential. Including their perspectives first allows for triangulation of the data and increased reliability. Triangulation allows to understand whether victims benefit from perpetrators participating in programmes (Lilley-Walker et al., 2016). This brings us to the second research question: *Is treatment impact based on perpetrators' self-reports higher compared to perpetrators' treatment impact based on (ex-)partners' reports?*

H4: Perpetrators' self-reports would show a larger decrease of their own perpetration of emotional violence than (ex-)partner-reports about the perpetrator.

This study builds upon the research avenue promoted by the DAPHNE III IMPACT project of the European Union, focused on the improvement, monitoring, and evaluation of perpetrator programmes in Europe (Akoensi et al., 2012; Hamilton et al., 2012; Lilley-Walker et al., 2016). As a result, the European Network for Work With Perpetrators (WWP-EN, 2014) developed the IMPACT Monitoring Toolkit for European perpetrator programmes, from which this study draws its data.

An added benefit of the current study design is to analyse how dyadic disagreement and abuse change over time, potentially as a result of programme participation. By failing to differentiate between such partner and actor effects, IPV is unduly viewed as static rather than evolving (Kim et al., 2008) argue. Many dyadic studies on IPV focus on mutual violence, but (dis)agreement about relationship status and hopes for the relationship are understudied. This study aims to contribute to the body of research through a) using multilevel analysis comparing multi-site longitudinal evaluations b) comparing perpetrator and (ex-)partner reports and c) patterns of participation, non-participation, and attrition. A review of European perpetrator programmes has shown that rates and reasons of attrition are hardly registered (Lilley-Walker et al., 2016). This obscures the conclusions to be drawn about the effectiveness of perpetrator programmes (Lilley-Walker et al., 2016; Westmarland & Kelly, 2013). Appendix 2 contains analyses of attrition, comparing drop-outs and completers.

Figure 1. Conceptual framework of the effects of motivation and relationship disagreement on emotional violence.



2. Theory

2.1 Motivation

Existential versus functional motivation

In a systematic synthesis of qualitative studies about male IPV perpetrator's perspectives on intervention and change, a recurring distinction was observed between existential and functional motivations (McGinn et al., 2020). The issue of motivation is relevant, as lack of motivation is associated with higher risk of recidivism (Jewell & Wormith, 2010). Existential motivation refers to desiring to change because it is morally right, whereas functional motivation means wanting to change because it helps in getting what one wants and involves self-preservation. Functional reasons are instrumental, such as participating in a programme to minimise negative court verdicts, preventing one's partner from leaving, winning back one's (ex-)partner, or regaining access to children or avoiding custodial care proceedings. Existential reasons are intrinsic, such as wanting to be a better person or a better father (Stanley et al., 2012; Gray et al, 2016).

The question then arises which type of motivation is associated with higher treatment impact. Some scholars suggest that both types of motivation lead to perpetrator change. In joint treatment, for example, voluntarily motivated clients may have a spill-over effect on the motivation of court-mandated clients (Donovan & Griffiths, 2015). Qualitative research suggested that both existential and functional factors can increase the motivation to change, and that functional motivation may be sufficient in securing initial programme participation, but additional drivers are required for ongoing commitment (Stanley et al, 2016). Other qualitative interviews showed for instance, that (anticipated) losses (such as contact with partner or

children) motivated men to take steps to change (Hester et al., 2006). However, most qualitative studies in a systematic synthesis emphasised that a minimal level of existential motivation was required for treatment (McGinn et al., 2020). Overall, these studies lacked methodological rigour, unfortunately (McGinn et al., 2020).

Motivation is no panacea

Even though motivation appears to be an important mechanism within perpetrator programmes for change, it is not a panacea. Cognitive distortions, issues of emotional regulation, patriarchal social constructions, and low self-esteem are barriers to change (McGinn et al., 2020), regardless of motivation. Likewise, the ineffectiveness of programmes is explained by low working alliance between therapist and client, high attrition rates, limited treatment engagement, and low accountability (Santirso et al., 2020). Furthermore, results from a meta-analysis showed motivation to be a dynamic process requiring ongoing investment by professionals (Babcock et al., 2004). Moreover, encouraging existential motivation during treatment leads to a stronger reduction of abuse, as shown in a systematic review and meta-analysis of randomised controlled trials (Santirso et al., 2020).

In sum, three conclusions can be taken: 1) both existential and functional motivation seem to be able to exist simultaneously and lead to change; 2) but existential motivation shows larger effects than functional motivation and 3) investing in existential motivation is worthwhile in affecting change.

2.2 Relationship context

The importance of couple reports

Several scholars have propagated the use of (ex-)partner reports besides self-reports in perpetrator programmes to increase validity and reliability (Ginés et al., 2015; Lilley-Walker et al., 2016; Westmarland & Kelly, 2013). Perpetrators tend to under-report and trivialise their abusive behaviour and its impact (Henning & Holdford, 2006). Other scholars have stressed also the importance of studying relationships at the dyadic level rather than only at the individual level (Cook & Snyder, 2005; Robins et al., 2000; Stephenson et al., 2013). IPV research sometimes lacks linked partners (Westmarland & Kelly, 2013), a limitation which the current study seeks to overcome. By doing so, (ex-)partners' influences on individuals can be dissected from the individual's contribution to their own behaviour (Kim et al., 2008). Following this reasoning, not only the perpetrator's own hopes for the future of the relationship can influence their motivation and their actual change, but also whether their (ex-)partner (dis)agrees with this.

Mechanisms in dyads: from expectations, preferences and disagreement to IPV

In order to explicate the relation between motivation and behaviour within the IPV context, it is necessary to examine the interdependence between hopes, preferences, expectations and disagreement within romantic couples. One's own behaviour is strongly motivated by expectations and evaluations of other people's behaviour (Guerrero & Floyd, 2006). Expectations can be based on general norms, but become more specific with people who know each other. Expectations are learned and thus can be changed accordingly (Guerrero & Floyd, 2006). This is also the case in abusive intimate relationships; hostile behaviour becomes the expected behaviour,

although non-hostile behaviour still remains the preferred behaviour (Jaspaert, 2015). If preferences are not met, one can respond by avoiding, approaching positively or reciprocating (Jaspaert, 2015). If people's preferences are not met and they disagree about their preferences, this has negative consequences (Jaspaert, 2015). Indeed, a meta-analysis confirmed that relationship dissatisfaction is a significant risk factor for IPV perpetration, especially for men (Stith et al., 2008).

A cross-sectional study shows that the couple's preferences regarding the relationship predict IPV in different ways (Jaspaert, 2015). Two main concepts were examined in that study; preference discrepancy, which refers to a gap between the individual's preferences and their perception of reality, and preference incongruence, which refers to disagreement in preferences at the dyadic level. Preference discrepancy was found to predict emotional IPV perpetration indirectly through relationship satisfaction. For men, discrepancy directly predicted emotional IPV perpetration. Disagreement in a couple about their preferences indirectly predicted IPV perpetration, through relationship satisfaction. Compared to preference discrepancy, preference disagreement was poorer in explaining the variance in relationship satisfaction and IPV perpetration. Despite the study's robust theoretical framework and use of Structural Equation Modelling, it remains to be seen whether these insights can be translated to the context of perpetrator programmes. The couples were sampled from the researcher's personal networks, not from perpetrator programmes or victim services.

2.3 Research on programme participation, relationship context and motivation

Then, what happens once programme participation becomes part of the equation, producing new expectations and preferences within both partners? In perpetrator

programmes, special attention is required to facilitate the management of couples' expectations of programme participation (Chung et al., 2020). Not only because perpetrators expect victims to stay with them because they are making an effort to change (McGinn et al., 2019), but also because perpetrators expect to change victims' minds on separation, through programme participation (Gray et al., 2016). As a result, victims should be offered realistic expectations around perpetrators' change, including the possibility of no change or a shift in the type of abuse (Kelly & Westmarland 2015; Vlasis 2014). Expectation management is central to victim safety, so victims can act on behalf of their own and children's safety in deciding whether to maintain or dissolve the relationship.

Research remains scarce assessing programme participation, the relationship context and motivation simultaneously (e.g. Crane et al., 2013; Downes & Vall, 2020; Gray et al., 2016). Primary analysis of a subset of the current data, showed different response patterns between clients and victims regarding their relationship status (Downes & Vall, 2020). Over time, the proportion of men and women reporting a dissolved relationship increased. However, women indicated more often to be separated, whereas men indicated more often to be still together, at the same timepoints. These patterns were not further examined in that study.

The interdependency between motivation, relationship context and victim safety was also found in a qualitative study among male participants (Gray et al., 2016). The study found that relationship statuses often changed during the programme, a proportion of the perpetrators were motivated to participate to "win" back their (ex-)partner and victims had difficulty trusting their (ex-)partners' motivation to participate (Gray et al., 2016). Perpetrators' motivation was also found to diminish because of separation (Gray et al., 2016). Thus, when a perpetrator is mainly

motivated to participate in a programme to get back their (ex-)partner and it becomes apparent this is not going to happen, the motivation to change disappears. Conversely - as McGinn and colleagues (2020) asked - if the perpetrator does win back his partner, what happens to his motivation when he is back home? If the goal is reached, the motivation to remain non-violent loses its function.

The effect of the relationship context on motivation is worrisome, as abuse often does not end because the perpetrator changes. Violence ends either because the relationship ends, or it continues or transforms after separation (Walby & Allen, 2004). A third study among men in a court-ordered programme supported these results, finding men who did not accept the relationship status to be more abusive than men who accepted the relationship status (Crane et al., 2013). Although verbal aggression was high for the non-accepting group, it declined over time (Crane et al., 2013).

2.4 Implicit implications of programme participation

Attention to the relationship context is warranted for two reasons, relating to the assumption of behavioural change within perpetrator programmes. First, the hyper-focus of perpetrator programmes on behavioural change pushes safe separation to the background (McGinn et al., 2019). Safe separation entails that the (ex-)partner feels safe enough to decide to separate from their abusive (ex-)partner, without abusive repercussions, harassment or acrimony (Mayer, 2017). However, (ex-)partners experience various pressures to continue the relationship while perpetrators are in treatment. Some victims stay with their abuser hoping the treatment changes them (Austin & Dankwort, 1999), whereas others feel forced to stay with their abuser because he makes an effort to change (McGinn et al., 2019). This is in line with the

assumption of programmes that perpetrators *should* take responsibility for their abuse and such behaviour can be unlearned (Lilley-Walker et al., 2016). However, victims have also noted their partners' commitment to the treatment is "guided by a continuing relational power dynamic" (Gray et al., 2016, p. 7).

A second concern highlights how perpetrator programmes might promote apparent rather than real change, causing victims to continue to live in unsafety. Perpetrator programmes can further teach perpetrators to conceal their abuse better, rather than to actually change (Donovan & Griffiths, 2015; McGinn et al., 2019). Apparent change also is observed when solicitors recommend programme participation to benefit the court sentence (McGinn et al., 2019; Westmarland & Kelly, 2013). This begs the question what programme evaluations are exactly measuring and monitoring.

Thus, taking into account the implicit effects of programme participation on victims and relationship dynamics has several positive effects. It allows for management of more realistic expectations around perpetrators change (Kelly & Westmarland 2015; Vlasis 2014). Also, it allows a more customised approach in managing grief, loss and acceptance related to separation which enables safe separation and safety planning as viable options (Mayer, 2017). As such, treatments can be more effective and have fewer adverse consequences for victims.

3. Method

3.1 Procedure

The questionnaire was administered at a maximum of five time-points. T0 was measured before the start of the programme, T1 at the beginning at the programme, T2 in the middle, T3 towards the end and T4 after the programme ended. Data

(n=503) were collected through nine domestic violence perpetrator programmes from four European countries. The organisations worked with perpetrators of domestic violence and collaborated with victim services mainly from other departments. Organisations' programmes varied in duration and it was not feasible for every programme to administer the survey at all timepoints. Some service providers experienced it was too burdensome for participants to complete five questionnaires. For others it depended on whether contact was possible with perpetrators and their (ex-)partners before programme commencement and after completion (Downes & Vall, 2020).

3.1 Participants

Participants were male perpetrators who attended domestic violence perpetrator programmes, and their female (ex-)partners who did not attend the programmes. The sampling procedure was non-random. (Ex-)Partner participation was not always possible due to legal or safety reasons. Table 1 and 2 show the participation rates varied across timepoints for clients and (ex-)partners, which indicates many unique entries. Tables 3 and 4 show that the participation rates for clients and (ex-)partners differ strongly per organisation.

Table 1. Participation rates for clients across all timepoints.

	C0	C1	C2	C3	C4	
C0	79					
C1	7	333				
C2	36	29	69			
C3	34	78	53	123		
C4	5	0	8	8	8	
Total	79	333	69	123	8	423

Table 2. Participation rates for (ex-)partners across all timepoints.

	P0	P1	P2	P3	P4
P0	34				

P1	5	38				
P2	16	22	41			
P3	6	16	23	29		
P4	2	0	4	3	4	
Total	34	38	41	29	4	80

Table **Error! No text of specified style in document.**3. Participation of male clients per organisation, across all timepoints.

	T0	T1	T2	T3	T4	Total
ORG.BU1	0	2	6	5	3	6
ORG.CR1	0	6	6	6	0	6
ORG.IT1	19	3	16	18	5	20
ORG.IT2	6	4	4	3	0	6
ORG.UK1	0	49	17	9	0	51
ORG.UK2	1	0	0	0	0	1
ORG.UK3	8	269	0	66	0	284
ORG.UK4	44	0	20	16	0	48
ORG.UK5	1	0	0	0	0	1
Total	79	333	69	123	8	423

Table 4. Participation of female (ex-)partners per organisation, across all timepoints.

	T0	T1	T2	T3	T4	Total
ORG.BU1	1	2	6	5	2	6
ORG.CR1	0	6	6	6	0	6
ORG.IT1	13	0	9	1	2	14
ORG.IT2	6	4	4	2	0	6
ORG.UK1	0	24	9	7	0	25
ORG.UK2	1	0	0	0	0	1
ORG.UK3	0	0	0	0	0	0
ORG.UK4	13	2	7	8	0	22
ORG.UK5	0	0	0	0	0	0
Total	34	38	41	29	4	80

To facilitate data-analysis, maintain a large enough sample size and ease interpretation, T0, T1 and T2 were pooled together, T3 was maintained and T4 was excluded. Many entries at T0, T1 and T2 were unique, whereby continuous or consecutive participation rates were low. For those who participated in at least two of these three measurements, little change was observed, see Appendix 2. Therefore, in

case of repeated entries at either T0, T1 or T2, the first entries were taken to ease comparison with others who only have unique entries.

The low rates of repeated participation had several causes. As the seven organisations worked within different contexts, such as jurisdictions, organisational structures, methods and trouble maintaining clients, their ways of collecting data varied strongly. Some organisations intentionally only collected data at two out of five timepoints, whereas others administered the questionnaire three times. Complicating matters even more, the organisations varied in which two or three timepoints they administered. Deciding to conservatively include only T1 and T3, would imply a large loss of (ex-)partner data, as organisations also differed in the opportunity to include (ex-)partners. Thirteen (ex-)partners were omitted from the sample as their partner had left the programme or had not administered any questionnaires. Exclusion of T4 did not result in much loss of cases, seeing that all but one participated at T3 as well. Appendix 2 describes the differences between drop-outs and completers.

The remaining sample size was $n = 490$, consisting of 423 male clients (86.3%) and 67 female (ex-)partners (13.7%). Of the sample, 27.3% ($n = 134$) formed a couple and 72.7% ($n = 356$) participated solo. Two-thirds (66.1%) of the whole sample participated at only one time-point; almost 70% of male clients ($n=293$) and close to half of the female (ex-)partners ($n = 31$). The remaining 130 male clients had more than two measurements, but 115 of them participated at the right two time-points (either T0, T1 or T2, *and* T3). For (ex-)partners, 36 participated at more than two time-points, but only 21 of them participated at the right two time-points.

Table 5 shows the demographic variables for clients at baseline. Most clients were aged between 22 and 50, with the largest group aged 31-40. Clients aged 18-21

were not represented in the subset. Almost seven in ten clients were full-time employed and other categories were represented but to a lesser extent. Overall, more than half of the clients had an income with which they managed to pay for essentials, but either had nothing left over or managed to buy the occasional treat or save sometimes.

Regarding referral route, voluntary routes were most common. The Croatian and Bulgarian organisations indicated that all their clients came to the programme as result of a court order (Association NAIA, 2019; Dom Duga-Zagreb, 2019). There was no mandatory participation in the Italian context, although participation could be strongly advised by social services or judicial authorities (E. Gajotto, personal communication, February 11, 2020; Una Casa per l’Uomo & Gruppo R, 2019).

Table 5. Demographic statistics (age, referral route, employment and income) of male clients at baseline.

Variable	Baseline in case of T3 (n=115)	
	n	%
Age		
18-21	0	0
22-30	25	21.7
31-40	40	34.8
41-50	27	23.5
51-60	12	10.4
over 60	3	2.6
Missing	8	7
Employment		
Full-time	80	69.6
Part-time	8	7
Unemployed	11	9.6
Sickness leave/retirement	9	7.8
Something else	7	6.1
Income		
Struggling to pay for the essentials	19	16.5
Managing to pay for the essentials, no left over	30	26.1
Managing to buy the occasional treat or save sometimes	32	27.8
Managing regular treats and saving or holiday	18	15.7
Comfortably managing	15	13
High income	1	0.9

Referral route compressed (multiple responses possible)

Mandatory	27	23.5
Voluntary	64	55.7
Other	12	10.4
Missing	12	10.4

3.2 Instrument

Perpetrators reported on their own abusive behaviour and (ex-)partners reported on their perpetrators' abusive behaviour towards them. The questionnaire contained items about the frequency, severity, and forms of intimate partner violence. Severity was measured through the impact of the abuse on the (ex-)partner. Other questions covered biographical information, referral route, reasons for using violence, impact on the (ex-)partner motivation to be in the programme, changes since the programme started, number of police call-outs, feelings of safety, relationship status and hopes for the future of the relationship.

The instrument was a self-reported questionnaire, which took between 15 and 30 minutes filling out, either online or on paper. Work With Perpetrators European Network (WWP-EN) created the IMPACT Monitoring Toolkit to increase women's and children's safety, through developing 'tools and methodologies to harmonise and enhance the monitoring and evaluation of work with perpetrators across Europe' (WWP-EN, 2015). Development of the toolkit was initiated by the IMPACT project and financially support by the European Union's DAPHNE III programme to prevent and combat violence against women and domestic violence.

The IMPACT instrument was developed and adapted from different existing instruments by WWP-EN project partners (Jones, 2017) and reviewed by practitioners from WWP-EN (WWP-EN, 2014). Items about the frequency and forms of emotional and physical violence were adapted from the revised Conflict Tactics Scale (CTS2).

This scale is criticised for not being able to take into account the context, motives and consequences of the abuse, and therefore coercive control, which contributes to a seeming gender-symmetry in perpetration (Kimmel, 2002; Jones, 2017). Inclusion of the impact of the violence on victims is essential, which led to the development of the COHSAR survey. See Hester et al. (2010) and Donovan and Hester (2014) for a discussion of the validation of the COHSAR instrument. COHSAR is also able to measure domestic violence in same-sex and heterosexual relationships and includes biographical information, relationship information and motivation for the behaviour. The COHSAR survey was adapted to the IMPACT survey on several points: longitudinal rather than cross-sectional; only male-perpetrated and female victims' experiences of domestic violence in heterosexual relationships; perpetrators' relationship with children rather than their parenting of the children; impact items were asked once after all abuse items rather than impact after every form of abuse; two items about well-being opposed to none (Jones, 2017). Questions about the number of police call-outs, relationship status and hopes for the future of the relationship were added specifically for the IMPACT survey.

No documentation was found on how motivation for coming to the programme was included in the IMPACT Monitoring Toolkit. The questions regarding reasons for attending are not included in the CTS, CTS-2 or COHSAR. Aforementioned WWP-EN documentation contains no reference to the development of the programme motivation items and whether the items are validated or could be combined in a scale.

3.3 Measures

The descriptive biographical variables were *age, employment, income, referral route* and were measured at first contact. The predicting variables were *existential*

motivation to participate, couple participation and relationship context disagreement.

The dependent variable was *emotional violence*.

Time was measured originally as T0) before the start of the programme, T1) at the beginning at the programme, T2) in the middle, T3) towards the end and T4) after the programme ended. T0, T1 and T2 were pooled together as '**baseline**'. In case of multiple entries per participant at baseline, the first time point was taken. T3 was identified as **follow-up** and T4 as omitted. Time was included as a predictor in the multilevel analyses.

Referral route was measured by asking clients how they came to the programme at T0, T1 and T2. Clients could provide more than one answer. In case clients provided both voluntary and mandatory routes, they were categorised as mandatory referrals. The variable was divided in two categories. **Mandatory participation** is composed of categories such as *criminal courts; probation or civil courts*. **Voluntary participation** is identified by items such as poster/internet/other publicity; helpline; friends/family/ colleagues.

Existential motivation to be in the programme was asked at T0, T1 and T2 (baseline) through *What are your reasons for coming to the programme?* Multiple dummy-categories could be checked. Factor analysis and reliability testing reveal that existential motives could be grouped together best, see Appendix 3. Participants could score between 0 - 7 on the scale. The scale contains items such as *I want to be a better parent to my children* and *I want to stop using violence*.

Relationship status was asked at all timepoints to both perpetrators and (ex-)partners by asking *What is your relationship status with your partner/ex now?* Answer categories were recoded into Together ('that we will be together and living

together), Apart ('that this relationship will end' or 'I am in another relationship already') and Other ('I am not sure' or 'Something else – please say what').

Relationship hope was operationalised by asking perpetrators and (ex-)partners at all timepoints *What are your hopes for your relationship with this person in the future?* The variable was modified by recoding it into Together ('that we will be together and living together'), Apart ('that this relationship will end' or 'I am in another relationship already') and Other ('I am not sure' and 'something else').

Relationship context disagreement was operationalised through comparison of the perpetrator's and (ex-)partners answers regarding their current relationship status and their hopes for the future of the relationship at all timepoints. No agreement nor disagreement was observed among solo participants, who formed the reference group for the multilevel analyses. Disagreement was observed if couples disagreed on at least one of the two aspects. For current relationship status, agreement was operationalised if both partners provided the same answer. Agreement was also observed for a combination of *together and living together* and *together but living apart*, or a combination of *in the process of splitting up* and *the relationship has ended and we are living apart*. Disagreement is measured by all other combinations.

Regarding hopes for the future of the relationship, agreement was measured if both partners marked the same answer, whereas disagreement is measured if they provide a different answer.

Emotional violence. Intimate partner violence was measured through various forms of emotional violence, perpetrated by the client against his partner or most recent ex-partner within the past 12 months. Emotional violence was measured through twelve items, such as *threats to hurt partner/ex*. Participants were asked at all timepoints to indicate the frequency of the violence: 1) never, 2) sometimes or 3) often. Sum-scores

were created from the frequencies and averaged, forming a continuous outcome variable ranging from 1 to 3.

3.5 Data analysis procedure

Descriptive statistics

SPSS (version 25) was used for descriptive statistics and for the preparation of the data for multilevel modelling. Univariate and bivariate statistics were inspected for biographical variables and all predictors. Frequencies were compared for completers at baseline and follow-up. Appendix 3 provides comparisons between drop-outs and completers regarding all variables. Significant differences within categorical variables were tested with chi-square tests and differences in means among scale variables were assessed through independent samples t-tests. For ordinal variables with more than two categories, significant changes in time within clients were tested with the non-parametric McNemar test of marginal homogeneity. These tests were performed for relationship status and relationship hopes. Changes in time of the outcome mean frequency of emotional violence were tested through paired sample t-tests.

Multilevel modelling

After initial data exploration and description, preparations were made for multilevel modelling. Multilevel modelling was conducted through MLwiN, version 2.36. Before presenting the analytical strategy, a few caveats should be noted relating to the dependencies at several levels. Participants potentially shared the same group therapy sessions, and a proportion of participants are part of a romantic couple, causing dependencies of observations and individuals. The data originally came from

nine organisations; only seven administered at least two measurements and a remaining five contained clients who participated at both baseline and follow-up. Both amounts are too low to be added as a third-level unit. Instead, they are added as dummy variables (Rasbash et al., 2017; M. E. Timmerman, personal communication, June 24 2021). For reasons of simplicity, the couple was not treated as a level and (ex-)partners were deleted after disagreement was calculated for coupled clients. Male participants were required to have completed the survey at least at two occasions and completed T3. For female participants, there was no requirement of completing specific timepoints or a minimum number of timepoints. This is inherent to the fact that (ex-)partners may have to drop-out from measurements out of safety or relationship separation, caused (partially) by relationship disagreement or functional motivations. (Ex-)Partners were only required to be part of a dyad, forming a(n) (ex-)couple with the perpetrator.

A two-level hierarchical model assessed the effects of *relationship context disagreement* and *motivation on emotional violence*. First-level units were timepoints nested within individuals (second-level unit). The model was built up in several steps, combining suggestions from Hox (2002) and Leckie et al. (2016) for longitudinal data.

Checking assumptions of normality, linearity and homoscedasticity was done through inspection of residuals. However, testing multiple models implies inspection of residuals after including all relevant predictors and parameters (Hox, 2002). As this is quite scrutinous and impractical, suggestions by Hox (2002) are followed. First, the two residual terms were examined for the baseline model with a random intercept and fixed slope of time (Hox, 2002; Leckie et al., 2016). Then, various residuals in the final model were inspected for violation of assumptions.

Conventionally, level-one variables were first added to the model (relationship context disagreement), followed by level-two variables (existential motivation). It started with the null model with only a random intercept and without any explanatory predictors. Time was added as a fixed slope in Model 2. This can be considered to be the null-model for longitudinal multilevel analysis. In line with Hox (2002), the covariates are now added as fixed effects. In Model 3, relationship context disagreement was added, comparing solo clients to agreeing coupled clients and disagreeing coupled clients. Model 4 organisation as a subject-level dummies. Model 5 included existential motivation. Then, Model 6 allowed time to be a random slope, varying between subjects. In Model 7, subjects' slopes were allowed to be random depending on their category of relationship context disagreement. Then, Model 8 allowed random slopes for existential motivation. Finally, in Model 9, interaction effects were included between time and motivation; time and coupled agreement; and time and coupled disagreement.

4. Results

4.1 Descriptive statistics

At baseline, the mean score in existential motivation was 3.30, with a large standard deviation of 2.44. Motivation was not measured at follow-up. Tables 6 and 7 present the reduced relationship status categories and relationship hopes. The percentage of clients and (ex-)partners who were not in a relationship increased over time, although the increase was larger for (ex-)partners and changes were not significant. Regarding hopes, most clients and (ex-)partners hoped to be together. While the percentages of both clients and (ex-)partners hoping to be apart increased over time, the changes were insignificant.

Table 6. Clients' relationship status and hopes, recoded

	Baseline (n=115)		Follow-up (n=115)	
	n	%	n	%
Relationship status				
Together	58	50.4	54	47.0
Apart	47	40.9	54	47.0
Other	8	7	5	4.3
Missing	2	1.7	2	1.7
Relationship hopes				
Together	65	56.5	55	48.7
Apart	13	11.3	24	20.9
Other	35	30.4	34	29.6
Missing	2	1.7	2	1.7

Table 7. (Ex-)Partners' relationship status and hopes, recoded.

	Baseline (n=21)		T3 (n=24)	
	n	%	n	%
Relationship status				
Together	13	61.9	11	45.8
Apart	7	33.3	12	50
Other	1	4.8	1	4.2
Missing	0	0	0	0
Relationship hopes				
Together	10	47.6	6	25
Apart	3	14.3	5	20.8
Other	8	38.1	9	37.5
Missing			4	16.7

Among completers, there were 75 solo clients and 36 coupled clients. Of these, only 38 clients and 24 (ex-)partners participated follow-up. As there were only 36 (ex-)partners who responded to relationship status and hopes, relationship context disagreement was measured among 36 clients. Table 8 displays frequencies of relationship context disagreement over time, At baseline, 25 of them disagreed with their (ex-)partner and eleven agreed. At follow-up, this changed to 30 disagreeing and six agreeing. There were no significant changes over time, $X^2(2) = 1.925, p = .38$.

Table 8. Clients' relationship context disagreement at baseline and follow-up.

	Baseline (n=111)	Follow-up (n=111)
--	------------------	-------------------

	n	%	n	%
Solo	75	67.6	75	67.6
Disagree and coupled	25	22.5	30	27.0
Agree and coupled	11	9.9	6	5.4

To test the fourth hypothesis, the decrease in intimate partner violence over time was compared between clients and (ex-)partners. Table 9 presents mean scores on IPV for clients and (ex-)partners in general, thereby disregarding dependencies between dyads. Table 10 shows these scores for participants linked to each other as a couple.

T-tests for paired samples in Table 9 indicated that the changes in average frequency was statistically significant for clients' self-reported and (ex-)partner reported emotional violence. The differences in decrease of emotional violence was also significant between clients and (ex-)partners. These results imply rejection of the fourth hypothesis: *Perpetrators' self-reports show a larger decrease of their own perpetration of emotional violence than (ex-)partner-reports about the perpetrator.*

Table 9. Clients' mean frequency of emotional violence according clients and (ex-)partners.

	Mean (S.D.)	Mean paired difference (S.D.)	t	df	Sig. (2- tailed)
<i>Emotional violence</i>					
Clients baseline	1.25 (.29)				
Clients follow-up	1.11 (.18)	.13 (.32)	4.42	109	<.01
(Ex-)Partners baseline	1.58 (.50)	.41 (.43)	4.36	20	<.01
(Ex-)Partners follow-up	1.18 (.39)				
Difference between clients and (ex-) partners in decrease		-.27	-3.41	129	<.01

Table 10 presents different results when the dependence within couples was accounted for. The change in violence could be calculated for thirty participants from fifteen linked couples. Here too, clients reported a smaller decrease of emotional

violence than (ex-)partners, but there was no significant difference between clients and (ex-)partners. As such, the results do not support the fourth hypothesis.

Table 10. Mean frequency of emotional violence and differences between linked partners from 15 couples.

	Mean (S.D.)	<i>t</i>	<i>df</i>	<i>p</i>
Emotional violence				
Clients baseline	1.30 (.50)			
Clients follow-up	1.13 (.18)			
(Ex-)Partners baseline	1.51 (.50)			
(Ex-)Partners follow-up	1.31 (.42)			
Decrease clients	-.17 (.51)			
Decrease partners	-.40 (.45)			
Difference in decrease	.23 (.79)	1.13	14	.28

4.2 Multilevel analysis

Subject- and time-level effects on emotional violence

Table 11 shows the results of a multilevel analysis with fixed effects and random intercepts that vary at subject-level and time-level. Numbers are given with three decimal places as they are very small. Model 1 contained only a random intercept of 1.180 which was the mean frequency of emotional violence across all clients and all timepoints. The random part of the intercept was represented by the level-2 variance of .009 and level-1 variance of .049. The intraclass coefficient (ICC) indicated that 15.5% of the variance of the two emotional violence measurements was between clients and a remaining 84.5% was variance across timepoints within clients.

Time was added as a fixed slope in Model 2, indicating a decrease of .133. This was the mean change in emotional violence between timepoints. The ICC coefficients remain constant; the distribution of the variance between level 1 and level 2 has not changed. Improvement of Model 2 compared to Model 1 was tested with a likelihood ratio test. This indicated that the reduction of 34.28 in the deviance score was significant, $p < .001$. Adding the fixed slope for time to the model provided a better

data fit. A model including a random intercept and fixed slope of time is commonly used as the null-model in longitudinal multilevel analysis (Hox, 2002).

Next, residuals were inspected to check assumptions of multilevel analysis. Normality was inspected with Figure 2, which plots the standardised residuals against the normal scores at level-one. The scatterplot does not follow a straight line, indicating the residuals for this model were not normally distributed. Figure 3 displays the normal plots at level-two, which has a slightly more straight line, but still does not cross the x- and y-axis at 0. Transformation of the emotional violence by base-10 logarithm did not improve normality sufficiently. As a result, the untransformed outcome was maintained, while bearing in mind that non-normality affects the reliability, validity and generalisability of coefficients and parameters.

The normal plots indicated two notable cases, which may be outliers. Upon inspection these clients scored high at baseline for emotional violence; 2.273 and 3.000. Separate coefficients for these two outliers were not included in the model, because they were thought to be representative of the true parameters. Contrastingly, the lower values among the remaining sample were likely to be a symptom of under-reporting.

Figure 2. Normal scores against time-level standardised residuals, for emotional violence untransformed (left) and log-10 transformed (right).

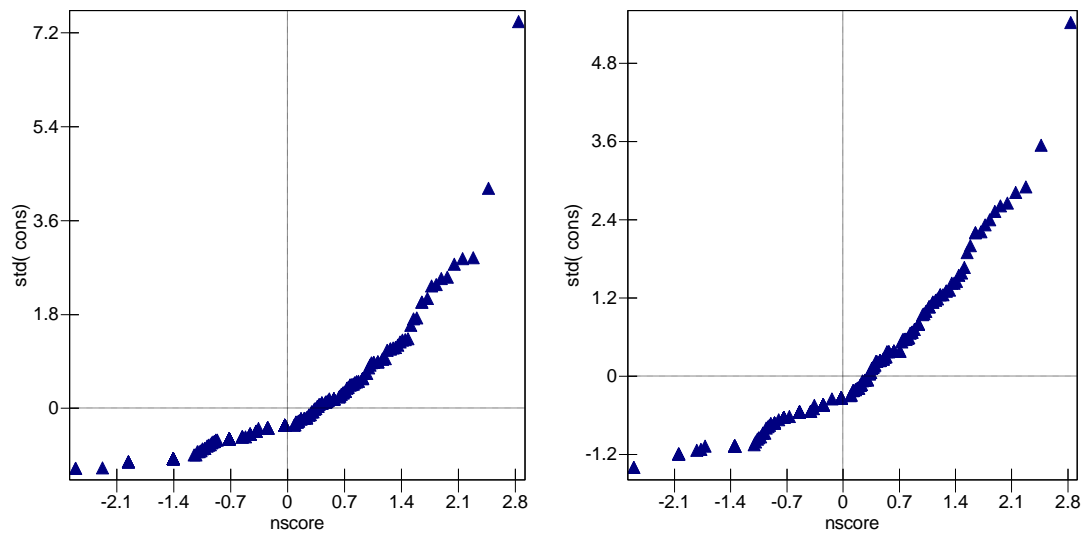
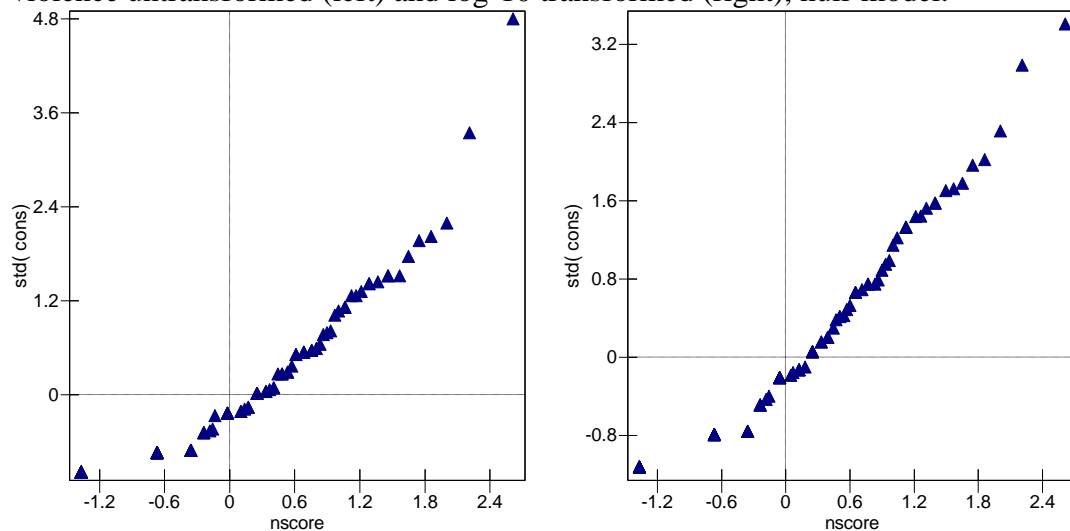


Figure 3. Normal scores against subject-level standardised residuals, for emotional violence untransformed (left) and log-10 transformed (right), null-model.



Inclusion of *relationship context disagreement* in Model 3 shows that coupled clients who disagreed, score significantly higher on emotional violence than solo clients. The variance between clients declines by .002, which translates to relationship context disagreement explaining 22.2% of the between-subject variance. The level-two ICC coefficient dropped from .155 to .13, slightly reducing the variation in emotional violence attributable to clients. The decrease in deviance was not significant, which means adding relationship context disagreement as a whole to the model did not significantly improve the fit of the data.

For Model 4, level-two dummies for organisations were added and the previous effects remained significant. All organisations differed significantly from the reference organisation. Inclusion of organisation explained 57.1% of the variation between clients and the decrease in deviance was significant.

Existential motivation was included in Model 5; making some of the differences between organisations less or not significant. The effect of coupled disagreement remained. Clients who reported more existential motives, scored significantly higher on emotional violence after adjusting for relationship context disagreement and organisation. As the level-two variance dropped to .000, existential motivation explained the full remaining between-subject variance. The reduction in deviance of 10.137 was significant; adding existential motivation to the model contributed to data fit.

Table 1. Results of multilevel analysis for emotional violence, fixed effects.

Model	M1: null model	M2: + time	M3: + relationship context disagreement	M4: + dummy organisations	M5: + existential motivation
	Estimate (S.E.)	Estimate (S.E.)	Estimate (S.E.)	Estimate (S.E.)	
Fixed part					
<i>Intercept and slope</i>					
Intercept	1.180 (.017)	1.246*** (.023)	1.211*** (.025)	0.955*** (.106)	1.028*** (.107)
Time (ref: baseline)		-.133*** (.030)	-.112*** (.030)	-.110*** (.030)	-.117*** (.032)
<i>Level-1 variables</i>					
Coupled and disagree (ref: solo)			.174*** (.039)	.202*** (.064)	.183*** (.066)
Coupled and agree (ref: solo)			.016 (.062)	-.016 (.073)	-.056 (.073)
<i>Level-2 variable</i>					
Organisation (ref. ORG.BU1)					
ORG.CR1				.388*** (.110)	.431*** (.108)
ORG.IT1				.283** (.109)	.237* (.097)
ORG.IT2				.283*** (.186)	.406* (.178)
ORG.UK1				.499* (.185)	.144 (.115)
ORG.UK3				.236* (.114)	.185 (.104)
ORG.UK4				.219* (.101)	.174 (.101)
Existential motivation					.020*** (.007)
Random part					
Level-two variance	.004 (.006)	.009 (.006)	.007 (.005)	.003 (.005)	.000 (.000)
Level-one variance	.058 (.008)	.049 (.007)	.047 (.007)	.047 (.006)	.048 (.005)
Level-two ICC	0.065	0.155	0.130	0.060	0.000
Level-one ICC	0.935	0.845	0.870	0.940	1.000
Deviance	12.919	-5.209	-17.951	-33.201	-44.144
Difference		18.128	12.742	15.250	10.943
Degrees of freedom		1	2	6	1

X^2 , p-value		<.001	<.001	<.001	<.001
Number of parameters	3	4	6	12	13
AIC	18.919	2.791	-5.951	-9.201	-18.144

* $p < .05$; ** $p < .01$; *** $p < .001$

The next step was to allow the slopes of the predictors to vary between subjects in each model, see Table 12. First, the effect of time was allowed to vary between subjects in Model 6. The ICC coefficients indicated that none of the variation was attributable to the time-level any more. The coefficients changed slightly, but the between-subjects variance increased from .000 to .065. This shows that clients varied more about their mean score at baseline, which was a more realistic estimation than with a fixed effect for time. The significant effect of coupled disagreement remained, but only ORG.CR1 differed significantly from ORG.BU1. Model 6 also contained two new parameters: the slope variance and the intercept-slope covariance. The negative covariance between intercept and slope, means that a more positive initial state was associated with a more negative, slope. Put differently, there was a ‘fanning in’ pattern; clients with a higher baseline score decreased more over time, whereas clients with a lower initial score decreased less over time. The likelihood ratio test shows Model 6 significantly fitted the data better.

A random slope for coupled agreeing clients was added in Model 7, but led to no reduction in unexplained variance. ORG.IT1 became significantly different from ORG.BU1 again. Allowing slopes to vary between groups did not change anything in the model, except for adding seven new (co)variance parameters. As such, emotional violence for solo clients, agreeing and disagreeing coupled clients can be predicted by the same slope.

Next, the slope of existential motivation was allowed to vary between subjects in Model 8. The reduction in deviance indicated that inclusion of this random slope did not contribute to a better data fit. The unexplained subject-level variance around the intercept decreased by 18.8%, but the unexplained variance for the effect of time was reduced only by 1%.

Finally, three interaction terms were included in Model 9; between time and coupled disagreement; time and coupled agreement; and time and existential motivation. Only the latter interaction was significant; which means that the decrease of emotional violence was stronger over time for clients with higher existential motivation. Inclusion of interaction terms removed any significant differences in clients' mean emotional violence between organisations. The random slope of time decreased from .096 to .088. This indicates that the degree of unexplained variance in the effect of time has decreased by 8.3% after inclusion of the three interaction terms. The deviance reduction of 10.806 was significant ($p = .013$), but comparison of the AIC coefficients indicated Model 6 was a better fitting model with the smallest AIC of -24.384. So, allowing the effects of existential motivation on emotional violence to vary between subjects and between time-points did not contribute enough to fit the data better.

Table 2. Results of multilevel analysis with predictors as random effects.

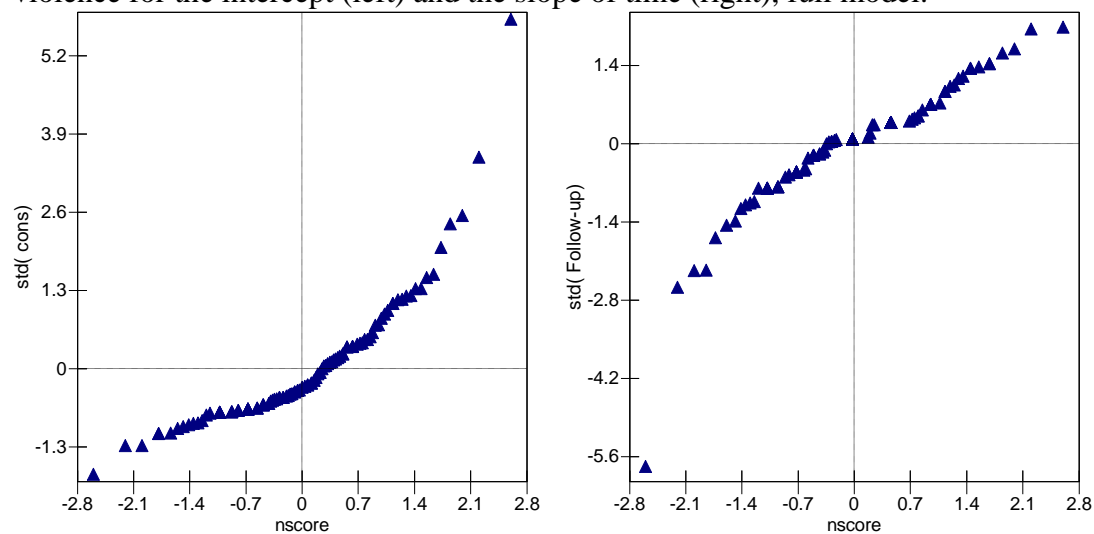
Model	M6: + time	M7: + relationship context disagreement	M8: + existential motivation	M9: + interactions
Fixed part	Estimate (S.E.)	Estimate (S.E.)	Estimate (S.E.)	
<i>Intercept and slope</i>				
Intercept	1.060*** (.106)	1.060*** (.106)	1.056*** (.099)	1.103*** (.097)
Time	-.117*** (.032)	-.117*** (.032)	-.118*** (.032)	-.101** (.033)
<i>Level-1 variables</i>				
<i>Relationship context disagreement (ref. solo)</i>				
Coupled and disagree	.168* (.068)	.168* (.068)	.151* (.066)	.190** (.070)
Coupled and agree	-.025 (.073)	-.025 (.073)	-.014 (.069)	-.026 (.080)
<i>Level-2 variables</i>				
Organisation (ref. ORG.BU1)				
ORG.CR1	.337*** (.101)	.337*** (.101)	.298*** (.094)	.292** (.092)
ORG.IT1	.192* (.096)	.192* (.096)	.206* (.087)	.154 (.089)
ORG.IT2	.332 (.171)	.332 (.171)	.327 (.163)	.258 (.166)
ORG.UK1	.146 (.112)	.146 (.112)	.136 (.105)	.084 (.105)
ORG.UK3	.157 (.101)	.157 (.101)	.158 (.094)	.110 (.094)
ORG.UK4	.132 (.099)	.132 (.099)	.123 (.094)	.069 (.094)
Existential motivation (grand mean)	.015* (.006)	.015* (.006)	.017* (.007)	.036*** (.010)
Time*coupled and disagree				-.210 (.100)
Time*coupled and agree				.008 (.112)
Time*existential motivation				-.032** (.012)
Random part				
cons/cons	.064 (.009)	.064 (.009)	.052 (.008)	.049 (.008)
Random slopes				
Time	.097 (.013)	.097 (.013)	.096 (.013)	.088 (.012)
Coupled and disagree			.000 (.000)	.000 (.000)

Coupled and agree		.000 (.000)	.000 (.000)	.000 (.000)
Existential motivation			.001 (.001)	.001 (.001)
Covariances*intercept				
Follow-up	-.064 (.010)	-.064 (.010)	-0.060 (.009)	-.056 (.009)
Coupled and disagree		.000 (.000)	.000 (.000)	.000 (.000)
Coupled and agree		.000 (.000)	.000 (.000)	.000 (.000)
Existential motivation			.004 (.001)	.004 (.001)
Remaining covariances				
Coupled and disagree*Time		.000 (.000)	.000 (.000)	.000 (.000)
Coupled and agree*Time		.000 (.000)	.000 (.000)	.000 (.000)
Coupled and agree* Coupled and disagree		.000 (.000)	.000 (.000)	.000 (.000)
Existential motivation*Time			-.005 (.002)	-.004 (.002)
Existential motivation* Coupled and disagree			-.000 (.000)	-.000 (.000)
Existential motivation* Coupled and disagree			-.000 (.000)	-.000 (.000)
Level-one variance	.000 (.000)	.000 (.000)	.000 (.000)	.000 (.000)
Level-two ICC	1.000	1.000	1.000	1.000
Level-one ICC	0.000	0.000	0.000	0.000
Deviance	-54.384	-54.384	-63.108	-73.914
Difference	10.240	0.000	8.724	10.806
Degrees of freedom	2	7	5	3
X^2 , p-value	.006	1.000	.120	.013
Number of parameters	15	22	27	30
AIC	-24.384	-10.384	-9.108	-13.914

* $p < .05$; ** $p < .01$; *** $p < .001$.

Lastly, diagnostics and assumptions were inspected for the better fitting Model 6. In Figure 5, level-two standardised residuals were plotted against their normal scores, both for the random intercept and the random slope of time. The residuals did not follow a straight line and the assumption of normality was still not met, as expected. These violations may result in incorrect estimates and inferences.

Figure 4. Normal scores against subject-level standardised residuals of emotional violence for the intercept (left) and the slope of time (right), full model.



Homoscedasticity assumptions were inspected through plots for standardised residuals against the fixed part of the model. Figure 6 shows relatively even spreads of the residuals along the y-axis, indicating homoscedasticity. There was some random scatter in both plots, but they did not cause a non-random shape or curvature. This also indicates that the data were predicted well by a linear relationship and the assumption of linearity was met. Figure 7 displays caterpillar plots where level-two random effects were plotted for all 111 clients, ranked by size. Most residuals' error bars did not cross the horizontal axis and thus differed significantly from the overall mean of emotional violence. There were many outliers, based on the mean score on emotional violence.

Figure 5. Subject-level residuals for emotional violence (standardised residuals against fixed part prediction), full model.

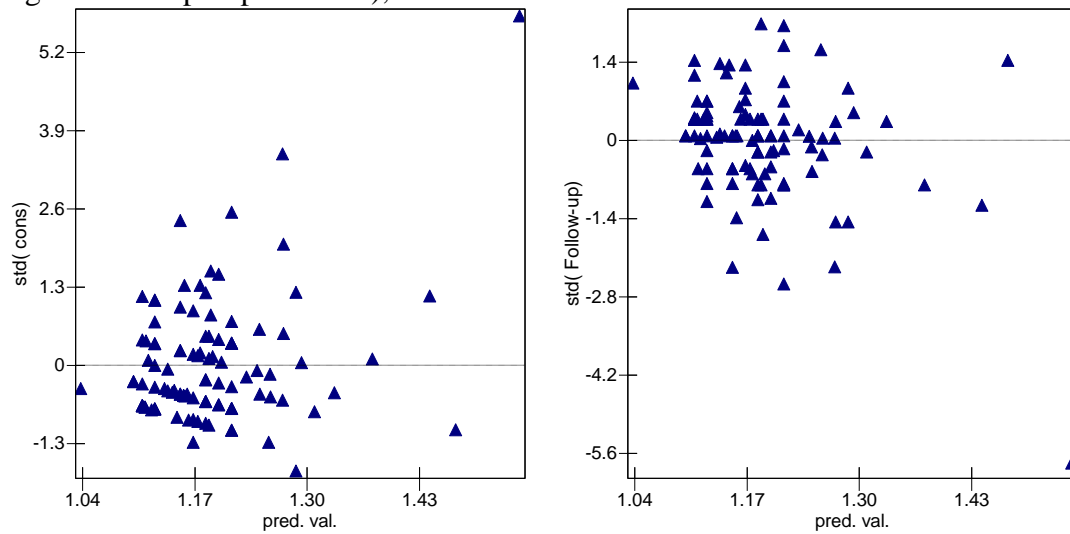
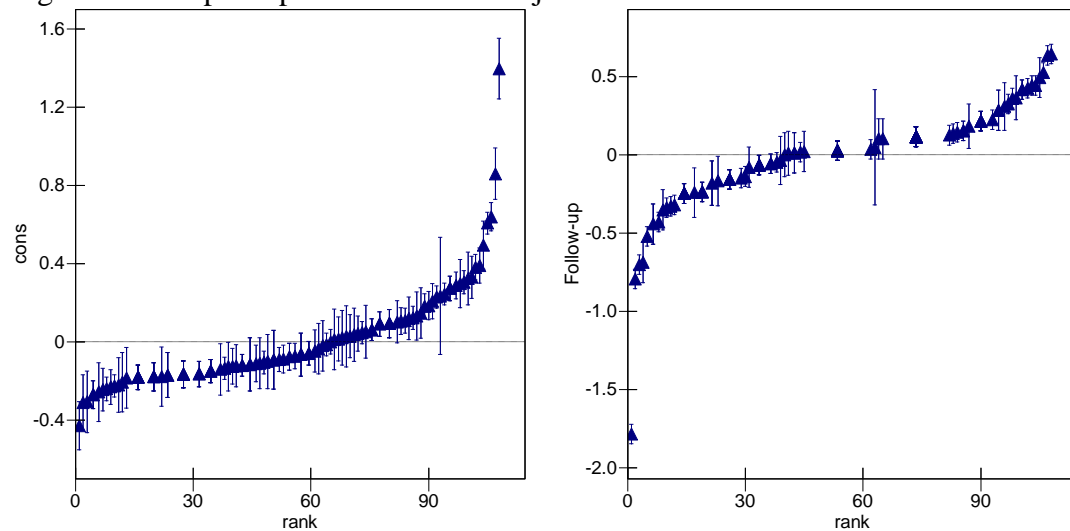


Figure 6. Caterpillar plots for ranked subject-level residuals.



Summary findings

In conclusion, the models with fixed effects (1 - 5) showed significant differences between clients in mean frequency of emotional violence, depending on their relationship context disagreement and existential motivation. Models 6 - 9 indicated that there were considerable differences between clients in the rate at which their emotional violence declined. The fixed effects pointed towards significant differences in baseline scores, based on relationship context disagreement, organisation and

motivation. Comparison of AIC coefficients pointed out that Model 6 was the best fitting model with fewest parameters. This means emotional violence significantly decreased over time after programme participation, for different categories of relationship context disagreement, organisation and levels of existential motivation. Model 9 indicated support for H1: the rate at which emotional violence decreased was stronger for clients who reported more existential motives. Results did not support H2, where the rate at which emotional violence decreased was thought to be steeper for agreeing coupled clients than solo clients. H3 was not supported either; emotional violence did not decline more over time for solo clients than coupled disagreeing clients.

5. Discussion and conclusion

In this study, the links between motivation, relationship context and intimate partner violence were examined among male clients participating in European perpetrator programmes through multilevel analyses. This study contributes to the existing literature through use of multi-site multi-country evaluations (Scambor et al., 2014), linked partner data (Westmarland & Kelly, 2013) and attrition analysis (Lilley-Walker et al., 2012), which are necessary but scarce within studies on the effectiveness of perpetrator programmes. Despite the setbacks of attrition, the study also contributes to the field through use of multilevel analysis and factor analysis for testing presumed classes of motivation. Also, the way in which motivation and the relationship context affect programme participation and IPV remain understudied so far (e.g. Crane et al., 2013; Downes & Vall, 2020; Gray et al., 2016).

Mapping these interdependencies is thought to be important for assessing programmes' effectiveness, which is not limited to the reported change in violence. Programme effectiveness extends to how programmes may unintentionally decrease

victim safety. Sometimes perpetrators participate, not because they want to change, but for reasons of self-enrichment (such as winning back their partner or avoiding a criminal sentence), which increases victim unsafety. Also, whereas couple participation can often increase treatment impact, the content of the relationship needs to be assessed. Clients may expect their partner to stay with them, because they participate in a programme. Disagreement about the wishes regarding the relationship may then lead to an increased risk of IPV and safe separation becomes less viable. In this case, measuring self-reported motivation alone is not sufficient. This is overcome by including the effects of the relationship context in the current study.

As such, the first expectation was that clients with higher existential motivation would report a stronger decrease of emotional violence over time (H1). However, it was also expected that couple participation would lead to a stronger decrease in emotional violence than solo participation, if the couple agreed about the relationship status and hopes for the future of the relationship (H2). Conversely, solo participation was expected to result in a stronger decrease of emotional violence compared to couple participation, if the couple disagreed about the relationship context (H3). Lastly, it matters whether clients or victims report programme effectiveness. It was expected that clients would report a stronger decrease of emotional violence than (ex-)partners (H4).

Results from multilevel analyses and t-tests indicated support for H1 and rejection of the remaining hypotheses, although emotional violence significantly decreased over time after programme participation. Clients with higher existential motivation scored higher on emotional violence, controlling for the effects of time, relationship context disagreement and organisation. Clients with higher existential motivation also showed a significantly stronger decrease in emotional violence over

time, than clients with lower existential motivation. Further, relationship context disagreement occurred frequently and disagreeing coupled clients had significantly higher baseline scores on emotional violence than solo clients. There were no significant differences between agreeing coupled clients and solo clients. As such, only a fixed effect was found for relationship context disagreement, but no random effects. H2 and H3 are thus rejected. Against expectations, (ex-)partners reported a significantly stronger decrease in emotional violence than clients when comparing unlinked partners. When participants were compared who were linked as a couple ($k = 15$), (ex-)partners also reported a stronger decrease than clients, but the differences were insignificant. This may be explained by lack of power due to the small sample size. In either case, of linked or unlinked partners, H4 is not supported.

Interpretation of the results requires some contextualisation. First, attrition was quite high; almost 70% of clients and close to 50% of (ex-)partners participated at one time-point. Results are therefore confined to clients who are more likely to complete the follow-up measurement. Second, organisations varied in when and how often they administered the survey, which may contribute to the differences in outcomes between organisations. This variation was already noted while developing the IMPACT Toolkit in 2015 (Ginés et al.), but apparently it has not improved since. Practitioners note that the IMPACT Toolkit is not easy to administer consistently and poses a high risk of drop-out (Hester et al., 2019). Third, four time-points were pooled together to a pretest posttest design, although multilevel analysis does not have many benefits in this case compared to regular ANCOVA (Hox, 2002). Fourth, the assumption of normality on which multilevel analysis relies was not met. Regarding the assumptions of linearity and homoscedasticity, there were only minor deviations. The absence of normality could impact the reliability, validity and generalisability of

the results negatively. These aspects are further compromised through absence of a comparison or control group (Lilley-Walker et al., 2016; Gondolf, 2004). Research implications may also not be applicable to lesbian, gay, transgender, queer, intersex (LGBTQI++) relationships (Morgan et al., 2019), because the study sample existed predominantly of clients in cis-heterosexual relationships. Lastly, the follow-up measurements were 'towards the end of the programme', so long term outcomes are not measured (Morgan et al., 2019).

Bearing these limitations in mind, the findings imply that emotional violence significantly declines over time, after treatment. Results however cannot prove the decrease occurs because of treatment. Results also imply that treatment impact is stronger for clients with higher existential motivation, who also have significantly higher baseline scores than clients with lower existential motivation. The higher frequency of violence may indicate a crisis as a need for change (Hester & Lilley, 2014). Also, treatment impact can be equally strong for solo clients and coupled clients, regardless of whether they (dis)agree with their (ex-)partner about the relationship status and/or hopes. Nevertheless, clients who disagree have significantly higher baseline scores on emotional violence than solo participants. The lack of differences in random effects between groups may be explained by a floor effect. This means that because the baseline score on emotional violence was very close to the minimum score of 1 to start with, a reduction is difficult to obtain. Results further found significant differences between (ex-)partner and clients reports about clients' use of intimate partner violence. This may be explained by under-reporting, where one party minimises the true frequency whereas the other party reports realistically. Most studies point to gender differences, where men tend to under-report their own use of violence, while women tend to under-report their own victimization (Chan,

2011). The results also found a ‘fanning in’ pattern, whereby clients with a higher baseline score on emotional violence decreased more over time than clients with a lower baseline score. Optimistically, this could point towards a stronger treatment impact for more severe perpetrators. However, it may also be the result of a floor effect; baseline scores were low and clients cannot decrease much more because they are close to the lowest range anyway.

To further this avenue of research, scholars may include measures of internal relationship discrepancy, which is observed when a client’s self-reported relationship status and hopes do not match. Future research may also specify which types of relationship context disagreement are a risk factor for IPV. Now contrasting categories were lumped together, although certain forms of disagreement may not be predictive of IPV. Likewise, further investigations with larger samples may help in distinguishing differential effects for existential versus functional motivations, as previous research indicated (Gray et al., 2016). The use of multilevel analysis could further be enhanced by incorporating organisation- or country-level effects as to discover factors contributing to treatment impact. Although multilevel analysis requires a limited number of variables especially with small samples (Tabachnick & Fidell, 2014), broader measures of success are essential in evaluating programme effectiveness. Reductions of severity, frequency and forms of violence or police call-outs are too limited (Westmarland & Kelly, 2013). Future research is therefore recommended to assess outcome variables from the IMPACT Toolkit such as impact of the violence on the victim and positive changes in the perpetrator since being on the programme, reported by both partners.

In conclusion, this study illustrated the complexity of studying clients and their (ex-)partners across various European programmes. The challenges in analysis that

resulted from drop-out, variation in administration and non-normality, also indicate directions for improvement and growth. Motivation and relationship context disagreement deserve further investigation under more robust research conditions. Also, this study has provided a multi-level set-up which can be used to explore other predictors and outcomes. Lastly, the reductions of emotional violence as reported by the (ex-)partner are important. This indicates that perpetrator programmes contribute to victim safety, because treatment impact is significant according to (ex-)partners, even when relationship disagreement is present. If the results are reliable, these are hopeful messages in regards of victim safety and safe separation.

Reference list

- 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(10), 1206–1225. <https://doi.org/10.1177/0306624X12468110>
- Association NAIA (2019). *WP 3: Testing of the operational protocol and elaboration of policy guidelines - D3.1 Reports on the testing*. Venice: Work With Perpetrators European Network - A.S.A.P. Project (unpublished material).
- Austin, J. B., & Dankwort, J. (1999). Standards for batterer programs: A review and analysis. *Journal of Interpersonal Violence*, 14(2), 152–168. <https://doi.org/10.1177/088626099014002004>
- 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(8), 1023-1053. <https://doi.org/10.1016/j.cpr.2002.07.001>.
- Campbell, J. C. (2002). Health consequences of intimate partner violence. *The Lancet*, 359(9314), 1331-1336. doi: 10.1016/S0140-6736(02)08336-8.

- Centre for Multilevel Modelling (n.d.). Errors and crashes. Retrieved from:
<http://www.bristol.ac.uk/cmm/software/support/support-faqs/errors.html>
- Chan, K. L. (2011). Gender differences in self-reports of intimate partner violence: A review. *Aggression and Violent Behavior, 16*(2), 167–175.
- Chung, D., Anderson, S., Green, D., & Vlasis, R. (2020). *Prioritising women's safety in Australian perpetrator interventions: The purpose and practices of partner contact*. Sydney: Anrows. Retrieved from
<https://www.anrows.org.au/publication/prioritising-womens-safety-in-australian-perpetrator-interventions-the-purpose-and-practices-of-partner-contact/>
- Corradi, C., Baldry, A., Buran, S., Kouta, C., Schröttle, M., & Stevkovic, L. (2018). Exploring the data on femicide across Europe. In Weil S., Corradi C., & Naudi, M. (Eds.), *Femicide across Europe: Theory, research and prevention* (pp. 93-166). Bristol: Bristol University Press. Retrieved from:
<http://www.jstor.org/stable/j.ctv8xnfq2.11>
- Crane, C. A., Hawes, S. W., Oberleitner, L. M. S., Mandel, D., & Easton, C. J. (2013). Relationship status acceptance, alcohol use, and the perpetration of verbal aggression among males mandated to treatment for intimate partner violence. *Journal of Interpersonal Violence, 28*(13), 2731–2748.
<https://doi.org/10.1177/0886260513487991>
- Dom Duga-Zagreb (2019). *WP 3: Testing of the operational protocol and elaboration of policy guidelines - D3.1 Reports on the testing*. Venice: Work With Perpetrators European Network - A.S.A.P. Project (unpublished material).
- Donovan, C., & Griffiths, S. (2015). Domestic violence and voluntary perpetrator programmes: Engaging men in the pre-commencement phase. *British Journal of Social Work, 45*(4), 1155-1171. <https://psycnet.apa.org/doi/10.1093/bjsw/bct182>

- Donovan, C., & Hester, M. (2014). *Domestic violence and sexuality: What's love got to do with it?* Bristol: Bristol University Press, Policy Press. Retrieved from <https://www.universitypressscholarship.com/view/10.1332/policypress/9781447307433.001.0001/upso-9781447307433>
- Downes, R. E., & Vall, B. (2020). *D5.2. Final research report A.S.A.P. Final report on the comparison between organisations that implemented the experimental operational protocol and the control group*. Berlin: European Network for Work with Perpetrators. Retrieved from [https://www.work-with-perpetrators.eu/fileadmin/WWP_Network/redakteure/Projects/ASAP/D5.2. Final_research_report_A.S.A.P.pdf](https://www.work-with-perpetrators.eu/fileadmin/WWP_Network/redakteure/Projects/ASAP/D5.2_Final_research_report_A.S.A.P.pdf)
- Elkin, M. (2019). *Domestic abuse victim characteristics, England and Wales: year ending March 2019*. London: Office for National Statistics. Retrieved from <https://www.ons.gov.uk/peoplepopulationandcommunity/crimeandjustice/articles/domesticabusevictimcharacteristicsenglandandwales/yearendingmarch2019>
- EIGE - European Institute for Gender Equality (2013). *Review of implementation of the Beijing Platform for Action in the EU member states: Violence against women - Victim support. Main findings*. Luxembourg: Publications Office of the European Union. Retrieved from <https://eige.europa.eu/sites/default/files/documents/Violence-against-Women-Victim-Support-Report.pdf>
- EIGE - European Institute for Gender Equality (2020). *Gender Equality Index 2020: Croatia*. Luxembourg: Publications Office of the European Union. doi:10.2839/29600.
- 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(2), 239-262. Retrieved from
<https://link.springer.com/article/10.1007/s11292-005-1179-0>
- FRA - European Union Agency for Fundamental Rights (2014). *Violence against women: An EU-wide survey - Main results*. Luxembourg: Publications Office of the European Union. Retrieved from
<https://fra.europa.eu/en/publication/2014/violence-against-women-eu-wide-survey-main-results-report>
- Ginés, C., Geldschläger, H., Nax, D., & Ponce, Á. (2015). European perpetrator programmes: A survey on day-to-day outcome measurement. *Studia Humanistyczne AGH*, 14(2), 33-52. doi:14.33.10.7494/human.2015.14.2.33.
- Gray, R., Broady, T., Gaffney, I., Lewis, P., Mokany, T., & O'Neill, B. (2016). 'I'm working towards getting back together': Client accounts of motivation related to relationship status in men's behaviour change programmes in New South Wales, Australia. *Child Abuse Review*, 25(3), 171–182. Retrieved from
<https://search.informit.org/doi/abs/10.3316/informit.167918781204495>
- Hamby, S. & Turner, H. (2013). Measuring teen dating violence in males and females: Insights from the national survey of children's exposure to violence. *Psychology of Violence*, 3, 323–339. doi: 10.1037/a0029706.
- 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(10), 1189–1205.
<https://doi.org/10.1177/0306624X12469506>
- Henning, K. & Holdford, R. (2006). Minimization, denial, and victim blaming by batterers how much does the truth matter? *Criminal Justice and Behavior*, 33, 110–130. <https://doi.org/10.1177%2F0093854805282322>

- Hester, M., Donovan, C., & Fahmy, E. (2010). Feminist epistemology and the politics of method: Surveying same sex domestic violence. *International Journal of Social Research Methodology*, 13(3), 251–263.
- Hester, M., Eisenstadt, N., Jones, C., Magnus, L., Morgan, K., & Bates, L. (2019). *Evaluation of year 2 of the Drive Project – A pilot to address high risk perpetrators of domestic abuse*. Bristol: Centre for Gender & Violence Research, University of Bristol. Retrieved from <https://driveproject.org.uk/wp-content/uploads/2020/01/Drive-Evaluation-Report-Final.pdf>
- Hester, M., & Lilley, S. J. (2014). *Domestic and sexual violence perpetrator programmes: Article 16 of the Istanbul convention. A collection of papers on the council of Europe convention on preventing and combating violence against women and domestic violence*. Strasbourg: Council of Europe. Retrieved from <https://edoc.coe.int/en/violence-against-women/7144-domestic-and-sexual-violence-perpetrator-programmes-article-16-of-the-istanbul-convention.html>
- Hester, M., Westmarland, N., Gangol, G., Wilkinson, M., O’Kelly, C., Kent, A., Diamond, A. (2006). *Domestic violence perpetrators: Identifying needs to inform early intervention*. Bristol: Bristol University, Northern Rock Foundation, Home Office. Retrieved from <https://equation.org.uk/wp-content/uploads/2016/02/EQ-LIB-135.pdf>
- Ivanova, S. (2016). *National study on domestic and gender-based violence - Analytical report*. Sofia: Partners Bulgaria Foundation. Retrieved from <http://old.csd.bg/fileSrc.php?id=23125>
- Jaspaert, E. (2015). *A dyadic analysis of the role of preference disconfirmation in the explanation of intimate partner violence* [Doctoral dissertation, University Leuven]. Retrieved from <https://lirias.kuleuven.be/1909489>

- Jewell, L. M., & Wormith, J. S. (2010). Variables associated with attrition from domestic violence treatment programs targeting male batterers: A meta-analysis. *Criminal Justice and Behavior*, 37(10), 1086–1113.
<https://doi.org/10.1177/0093854810376815>
- Jones, C. (2017). *Behind the IMPACT Toolkit* [PowerPoint slides]. Berlin: Work With Perpetrators European Network. Retrieved from: https://www.work-with-perpetrators.eu/fileadmin/WWP_Network/redakteure/Webinare/Behind_Impact/Webinar_Behind_The_IMPACT_Toolkit_FINAL_2017.pdf
- Kelly, L., & Westmarland, N. (2015). *Domestic violence perpetrator programmes: Steps towards change. Project Mirabal final report*. London and Durham: London Metropolitan University and Durham University. Retrieved from <https://repository.londonmet.ac.uk/1458/1/ProjectMirabalfinalreport.pdf>
- Kimmel, M. S. (2002). 'Gender symmetry' in domestic violence: A substantive and methodological research review. *Violence Against Women*, 8(11), 1332–1363.
<https://doi.org/10.1177%2F107780102237407>
- Kitzmann, K. M., Gaylord, N. K., Holt, A. R., & Kenny, E. D. (2003). Child witnesses to domestic violence: A meta-analytic review. *Journal of Consulting and Clinical Psychology*, 71(2), 339–352. <https://doi.org/10.1037/0022-006X.71.2.339>
- Leckie, G., Morris, T., & Steele, F. (2016). Multilevel modelling of repeated measures data - MLwiN practical. LEMMA vle module 15, 1-80. Retrieved from: <http://www.bristol.ac.uk/cmm/learning/online-course/>.
- Lilley-Walker, S. J., Hester, M., & Turner, W. (2016). Evaluation of European domestic violence perpetrator programmes: Toward a model for designing and reporting evaluations related to perpetrator treatment interventions. *International*

Journal of Offender Therapy and Comparative Criminology, 62(4), 868-884.

<https://doi.org/10.1177%2F0306624X16673853>

Mayer, A. G. K. (2017). Intervening with couples experiencing domestic violence: Development of a systemic framework. *Australian and New Zealand Journal of Family Therapy*, 38(2), 244-255. Retrieved from

<https://onlinelibrary.wiley.com/doi/abs/10.1002/anzf.1217>

McGinn, T., McColgan, M., & Taylor, B. (2020). Male IPV perpetrator's perspectives on intervention and change: A systematic synthesis of qualitative studies.

Trauma, Violence, & Abuse, 21(1), 97–112.

<https://doi.org/10.1177/1524838017742167>

McGinn, T., Taylor, B., & McColgan, M. (2019). A qualitative study of the perspectives of domestic violence survivors on behavior change programs with perpetrators. *Journal of Interpersonal Violence*, 1-27.

<https://doi.org/10.1177/0886260519855663>

Morgan, S. A., McCausland, B., & Parkes, J. (2019). Baseline characteristics and outcomes of the main perpetrator programme within the Hampshire Domestic Abuse Prevention Partnership, UK: A mixed methods study. *PloS One*, 14(7), e0218408. <https://doi.org/10.1371/journal.pone.0218408>

Morgan, K., Williamson, E., Hester, M., Jones, S., & Feder, G. (2014). Asking men about domestic violence and abuse in a family medicine context: Help seeking and views on the general practitioner role. *Aggression and Violent Behavior*, 19,

637-642. <https://doi.org/10.1016/j.avb.2014.09.008>

Oliver, R., Alexander, B., Roe, S., & Wlasny, M. (2019). *The economic and social costs of domestic abuse*. London: UK Home Office. Retrieved from

https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/918897/horr107.pdf

Rasbash, J., Steele, F., Browne, W., & Goldstein, H. (2017). A user's guide to MLwiN. Version 3.01. Bristol: Centre of Multilevel Modelling, University of Bristol. Retrieved from

<https://www.bristol.ac.uk/cmm/media/software/mlwin/downloads/manuals/3-01/manual-web.pdf>

Scambor, C., Wojnicka, K., & Scambor, E. (2014). *Possibilities for multisite/multi-country European evaluation studies on domestic violence perpetrator programmes. Working paper 3 from the Daphne III project "IMPACT: Evaluation of European Perpetrator Programmes"*. Berlin: Work With

Perpetrators European Network. Retrieved from: www.impact.work-with-perpetrators.eu/fileadmin/WWP_Network/redakteure/IMPACT/Daphne_III_Impact_-_Working_paper_3_-_Possibilities_for_multisite_multicountry_European_evaluation_studies_on_domestic_violence_perpetrator_programmes.pdf

Santirso, F. A., Gilchrist, G., Lila, M., & Gracia, E. (2020). Motivational strategies in interventions for intimate partner violence offenders: A systematic review and meta-analysis of randomized controlled trials. *Psychosocial Intervention*, 29(3), 175-190. Retrieved from <https://journals.copmadrid.org/pi/art/pi2020a13>

Siegel, J. P. (2013). Breaking the links in intergenerational violence: An emotional regulation perspective. *Family process*, 52(2), 163-178.

<https://doi.org/10.1111/famp.12023>

- Stanley, N., Graham-Kevan, N., & Borthwick, R. (2012). Fathers and domestic violence: Building motivation for change through perpetrator programmes. *Child Abuse Review*, 21(4), 264-274. <https://psycnet.apa.org/doi/10.1002/car.2222>
- Stark, E. (2007). *Coercive control: How men entrap women in personal life*. New York: Oxford University Press.
- Thornberry, T. P., Knight, K. E., & Lovegrove, P. J. (2012). Does maltreatment beget maltreatment? A systematic review of the intergenerational literature. *Trauma, Violence & Abuse*, 13(3), 135–152. <https://doi.org/10.1177/1524838012447697>
- Tobias, S., & Carlson, J.E. (1969). Brief report: Bartlett's test of sphericity and chance findings in factor analysis. *Multivariate Behavioral Research*, 4(3), 375-77. doi: 10.1207/s15327906mbr0403_8
- Videva, D. (2013). *Work with Perpetrators – European Network. National report Bulgaria*. Berlin: Work with Perpetrators – European Network. Retrieved from: http://www.work-with-perpetrators.eu/fileadmin/WWP_Network/redakteure/Reports_2013/National_report_Bulgaria.pdf.
- Una Casa per l’Uomo & Gruppo R (2019). *WP 3: Testing of the operational protocol and elaboration of policy guidelines - D3.1 Reports on the testing*. Venice: Work With Perpetrators European Network - A.S.A.P. Project (unpublished material).
- Vlais, R. (2014). *Domestic violence perpetrator programs: Education, therapy, support, accountability or struggle?* Victoria: No To Violence. Retrieved from <https://graddipfdr20161.pbworks.com/f/Domestic%20Violence%20Perpetrator%20Programs.pdf>
- Walby, S. (2005). Improving the statistics on violence against women. *Statistical Journal of the United Nations Economic Commission for Europe*, 22(3-4), 193-

216. Retrieved from <https://www.un.org/womenwatch/daw/egm/vaw-stat-2005/docs/expert-papers/walby.pdf>
- Walby, S., & Allen, J. (2004). *Domestic violence, sexual assault and stalking: Findings from the British Crime Survey*. London: UK Home Office. Retrieved from <https://nomsintranet.org.uk/roh/official-documents/HomeOfficeResearchStudy276.pdf>
- Widom, C. S. (1989). Child abuse, neglect, and adult behavior: Research design and findings on criminality, violence, and child abuse. *American Journal of Orthopsychiatry*, 59(3), 355-367. <https://content.apa.org/doi/10.1111/j.1939-0025.1989.tb01671.x>
- Westmarland, N., & Kelly, L. (2013). Why extending measurements of 'success' in domestic violence perpetrator programmes matters for social work. *The British Journal of Social Work*, 43(6), 1092-1110. Retrieved from <https://academic.oup.com/bjsw/article/43/6/1092/1653332>
- Woodlock, D. (2017). The abuse of technology in domestic violence and stalking. *Violence against Women*, 23(5), 584-602. <https://doi.org/10.1177%2F1077801216646277>
- Work With Perpetrators European Network (2014). *Impact Toolkit Guidelines 2014*. Berlin: Work With Perpetrators European Network. Retrieved from: https://www.work-with-perpetrators.eu/fileadmin/WWP_Network/redakteure/IMPACT/PDF_Toolkit_rev_2015/English/IMPACT%20toolkit%20guidelines.pdf
- Work With Perpetrators European Network (2015). *Impact Toolkit working paper*. Retrieved from: [5](https://www.work-with-</p></div><div data-bbox=)

[perpetrators.eu/fileadmin/WWP_Network/redakteure/IMPACT/Working_Papers/
Impact_Working_Paper_1_Ext.pdf](http://perpetrators.eu/fileadmin/WWP_Network/redakteure/IMPACT/Working_Papers/Impact_Working_Paper_1_Ext.pdf)

World Health Organization. (2013). *Global and regional estimates of violence against women: prevalence and health effects of intimate partner violence and non-partner sexual violence*. Geneva: World Health Organization. Retrieved from <https://www.who.int/publications/i/item/9789241564625>

Appendix 1 - Intimate partner violence estimates

This section provides a more detailed background of estimates surrounding male on female IPV perpetration. First, global, European and country-specific prevalence rates are discussed, from general population statistics. Then, service-based statistics and contexts are discussed for victims (residential and non-residential care) and for perpetrators. For perpetrator programmes, their general content is described first, then I shape the European context of programmes, the amount of programmes in the countries under study and their obstacles.

1. Intimate partner violence victimisation rates

Statistics continue to show that IPV is prevalent on an international, continent and national level and has far-reaching adverse consequences for victims and their children. Estimates for lifetime victimisation vary from 27 per cent on a global level (World Health Organization, 2013), to between 22 and 25 per cent on a European level (FRA - European Union Agency for Fundamental Rights, 2014; World Health Organization, 2013). Focusing on more recent victimisation within Europe, between 2 and 6 per cent of the women were victimised in the past year (FRA, 2014).

For the countries under study, estimates for lifetimes prevalence range from 13 to 29 per cent and recent prevalence ranges from 2 to 19 per cent. Without the outlier of 19% from Bulgaria's national survey, estimates range from 2 to 6 per cent. In Bulgaria, 23% of women experienced IPV ever, versus 6% in the past year (FRA, 2014). Compared to a national survey, lifetime IPV is reported by 25% of women from the general population (and by 43% of Roma women). Recent IPV is reported by 19% of the women (and 34% of Roma women) (Ivanova, 2016). In Croatia lifetime prevalence of IPV is 13%, versus 3% in the past twelve months (FRA, 2014).

In 2009, a national survey in Croatia estimated that 31% of women experienced frequent domestic violence, and 44% experienced it occasionally. No distinction was made between intimate partners and other members of the household, nor were the types of violence reported (WAVE, 2012). No recent comparative national statistics currently exist for Croatia (European Institute for Gender Equality, 2020). Lifetime prevalence for Italian women was 19% according to FRA (2014), compared to 14% in a national sample by Istat (2015). Recent IPV was reported by 6% of the women in Italy to FRA (2014) and 2% to Istat (2015). Lastly, in the U.K. 29% of women reported lifetime IPV to FRA (2014), versus 5% in the past year. This last estimate is similar to the 5% of women reporting recent IPV victimisation to the Crime Survey for England and Wales (Elkin, 2019). Lifetime IPV was not recorded by this source.

The estimates above have several limitations that complicate comparison and extrapolation. First, operationalisation and methodology affect prevalence rates and gender patterns strongly (Hamby & Turner, 2013). Most studies only measure physical and sexual violence, leaving out emotional, digital, and economic abuse. Nonetheless, such forms are just as coercive and controlling towards victims and can severely harm them both short- and long-term (Stark, 2007; Woodlock, 2017). Because the lines between unkind acts versus emotional abuse are sometimes unclear, there is no consensus about the appropriate operationalisation among researchers (World Health Organization, 2013). This is a reason that prevalence rates underestimate the true figures.

Second, the acceptance of talking about domestic violence and which behaviours are considered domestic violence vary per cultural context. These factors also explain variations in reported IPV which therefore do not necessarily reflect the actual prevalence of the violence (FRA, 2014). Last, it should be taken into account that the

estimates come from a general population which often do not include victims who are illiterate, fleeing from the violence, staying in shelters and therefore these figures are an underestimation (Walby, 2005). In sum, operationalisation, cultural contexts, and sample types contribute to an underestimation of the prevalence of IPV.

2. Service provision to victims and perpetrators

Whereas IPV is prevalent in general populations, it is even more severe and frequent in facility- or service-based populations (Walby, 2005). As discussed in the main text, the consequences of IPV on victims and children are devastating, seeking help remains complicated and help is not always effective. Estimates surrounding victim support services and perpetrator programmes provide insight into the current state of affairs. Once families have come to the attention of support services, either residential or non-residential help can be set in motion. Although the field has developed much over the past decades, the estimates below illustrate that proper care is still lacking. This increases the risks for victims and children who are forced to stay in an unsafe situation with their perpetrator. European estimates of service provision are provided, and country estimates when available. Whereas the most recent estimates available are presented, some estimates may be out-dated and do not reflect the current situation.,

2.1 Residential care for victims

The Istanbul Convention on Preventing and Combating Violence against Women and Domestic Violence (2011), signed by 45 out of 46 European countries, specifies one family space through sheltered accommodation per 10.000 inhabitants. On a European level, there is a deficit in these facilities as calculated by Women Against Violence Europe (2019). Among the 28 E.U. member states, there is one family space

per 20.000 inhabitants, which means 51% of the spaces needed do not exist. In other words, for every victim provided with safe accommodation, one victim cannot be sheltered. In Bulgaria and the UK, NGO's run all women's shelters, while this happens in collaboration with the state in Croatia and Italy.

2.2 Non-residential care for victims

Counselling centres

24 out of 29 member states have counselling centres for women survivors of IPV (European Institute for Gender Equality, 2013). Counselling centres offer gender-specific non-residential support to women who are victims of IPV and their children. Support ranges from information, advice, counselling, advocacy, practical support, court accompaniment, outreach and prevention. In 2012, only eight member states, including Croatia and the U.K., met the Istanbul Convention's recommendation to provide at least one counselling service per 50.000 women. that the assessment of support services is incomplete, as national data collection and evaluation is not undertaken regularly (EIGE, 2013).

Emergency services

In 2013, only thirteen out of 29 member states (of which Croatia, Italy and the U.K.) offer specialised emergency services for women who are victim of IPV (EIGE, 2013). These services are provided by specialised staff in hospitals, social services teams and shelters, mainly organising accommodation or proactive support. Only in five member states (including Italy and the U.K.) mobile psychosocial support is provided. Here specialised support workers reach out by going to where the victim lives or stays.

2.3 Perpetrator programmes

What are perpetrator programmes?

Domestic violence perpetrator programmes are treatments focused on affecting behavioural change, requiring perpetrators to take responsibility for their abusive and presuming such abusive behaviour can be unlearned (Lilley-Walker et al., 2016).

Programmes often deploy a combination of individual and group therapy and attend to increasing men's awareness, understanding and recognition of their violence and the impact on their (ex-)partner(s) and/or child(ren). Through this, they "aim to increase empathy, accountability and motivation to change, and challenge gender stereotypes and hostile attitudes towards women". Programmes offer a large array of approaches; cognitive-behavioural models, psycho-educational models, systemic therapy, psychodynamic models or a combination of approaches (Lilley-Walker et al., 2016).

Perpetrator programmes generally know three referral routes (Hester & Lilley, 2014);

- 1) Voluntary programmes either through self-referrals (which can be motivated by a crisis, when the (ex-)partner has left or is threatening to leave, or when child contact is threatened or complicated), or through pressure or advice from professionals;
- 2) Criminal justice programmes (after prosecution, possible in or outside of prisons, both mandatory or voluntary);
- 3) Community programmes (through referral from the criminal justice system without prosecution).

2.4 European context of perpetrators programmes

A focus on European programmes is warranted, as most of the literature stems from the American context, whereas practices are not always directly transferable to

other contexts (Hamilton et al., 2012). For instance, in the United States programmes are mostly autonomous and work without the support of partner organisations, whereas “European perpetrator programmes are usually embedded within a dense collaborative network of practitioners with a strong governmental influence” (Hamilton et al., 2012). Whilst programmes may benefit from exchange in knowledge and practices in Europe, there is no homogeneous European approach to working with perpetrators. Legal frameworks, national regulation and inter-institutional cooperation vary per country, but also the political and ideological interests, cultural aspects and allocation of resources differ from one another (Downes & Vall, 2020; Ginés et al., 2015; Hamilton et al., 2012). This is reflected in perceptions of what is considered to be domestic violence (FRA, 2014), how it should be approached (punitive or rehabilitative), the level of approach (individual, group or systemic; victim or perpetrator), which organisations are responsible for preventing, ending and providing care after domestic violence, and decision-making on funding (Hamilton et al., 2012).

2.5 Availability and obstacles of programmes in countries under study

Although there is variation between the countries under study, they also share certain trends regarding the recent rise of perpetrator programmes offered and obstacles experienced. Despite recent advances in the number of programmes offered, they still remain quite scarce in the countries under study. In 2012 for instance, there were no Bulgarian perpetrator programmes and only one Italian programme existed (Hamilton et al., 2012). Conversely, the UK has the longest track record with programmes running since 25 years, of which eleven were counted in 2012 (Hamilton et al., 2012) and sixteen three years later (Ginés et al, 2015). In 2013, experts estimated three prison programmes to exist in Bulgaria (Videva, 2013), whereas in

2015 one perpetrator programme was reported (Ginés et al., 2015). Whereas there were no data available for Croatia in 2012 (Hamilton et al.), seven programmes were listed in 2015 (Ginés et al.). In 2016, approximately 25 Italian perpetrator projects existed providing different levels of service. This seems like a steep increase compared to 2015, when Ginés and colleagues collected responses from nine programmes. According to the authors this probably does not represent the true amount of existing programmes. On the other hand, programmes were only introduced there in 2009 and male violence against women was included in legislation between 2013 and 2015 (Pauncz, 2016), which may account for the increase.

Perpetrator programmes in the four countries under study, experience various obstacles affecting the degree to which they can provide proper service. In Italy, perpetrator programmes experience obstacles such as achieving recognition, securing funding, developing collaboration with victims services and developing evaluation of programmes (Pauncz, 2016). Issues with delivery of perpetrator programmes are also experienced in Croatia, with only two community programmes in 2016 (Vurusic, 2016). Although preliminary steps were taken in legislation towards perpetrators of domestic violence in Bulgaria, programmes are not recognised and funded sufficiently by the Department of Justice (Vurusic, 2016). Due to inadequate funding, an insufficient number of programmes exist and professionals are trained. Practically, perpetrators cannot be asked to travel more than 50 kilometers to attend treatment and thus judges cannot order treatment in regions with lacking services (Vurusic, 2016). Although perpetrator programmes appear most developed in the UK compared to the other three countries, issues in funding and accessibility are shared. From 2010 onwards, programmes also struggled with service provision due to loss of public sector funding, changes in government policy and commissioning. This has affected

voluntary programmes and small specialised services most (Respect, 2016), seemingly pushing towards a more punitive and generalised approach in working with perpetrators.

Appendix 2 - Attrition

Attrition in general

Attrition within perpetrator programmes pertains to non-completion of treatment and/or evaluation by clients. Several authors warn that significant treatment effects may be incorrectly found, when programme completers differ significantly from those who dropped out or were excluded in the pre-commencement phase (Babcock et al., 2004; Feder & Wilson, 2005). These scholars suggest that completers may have higher motivations or more fear towards criminal justice sanctions. Critique is noted that drop-out rates have unduly mainly been examined for court-mandated programmes (Donovans & Griffiths, 2015). Other factors are also associated with higher probability of attrition, such as “employment, age, income, education, marital status, race, referral source, previous domestic violence offences, criminal history, and alcohol and drug use”, according to a meta-analysis (Jewell & Wormith, 2010, p.1). But more importantly, attrition predicts recidivism of domestic violence (Jewell & Wormith, 2010; Kraus, in Lilley-Walker et al., 2016). Thus, both practitioners and researchers struggle with the implications of drop-out rates for the presumed effectiveness of perpetrator programmes (Westmarland & Kelly, 2013).

Therefore it is striking that in a review of sixty European programme evaluations, attrition rates and reasons for attrition were hardly reported (Lilley-Walker et al., 2016). Nor did many studies investigate attrition further by answering who drops out, when and why, although this is essential “to enable to enable inferences about statistical power and the ability to generalise findings to wider populations” (Lilley-Walker et al, 2016, p.11). The relation between attrition and administering evaluations among clients is a complex one. Case managers indicated that outcome measurement tools such as the IMPACT Toolkit, were difficult “to use consistently and posed a

high risk of disengagement”, especially regarding the precarious engagement of clients (Hester et al., 2019). Paradoxically, evaluations are less reliable because of attrition, but at the same time evaluations promote attrition.

Present study and problems with attrition

This study aims to improve on issues of attrition mentioned above, but the design has several shortcomings. First, although the IMPACT Toolkit is developed to overcome many issues encountered in previous evaluations, it does not register reasons for attrition. Second, it was not possible to investigate perpetrators who are excluded from participation. Third, this study draws conclusions about attrition regarding evaluation participation, which might differ from treatment participation. In turn, treatment adherence does not necessarily imply actual behavioural change (Lilley-Walker et al., 2016).

This study investigates the differences between completers and drop-outs. Completers are understood as perpetrators who participated at at least baseline and follow-up. Drop-outs are participants who did not complete T3. First, original participation rates are examined across organisations. Then, the pooling procedure is explained. This is followed by comparisons between drop-outs and completers regarding biographical variables, couple participation, relationship context disagreement, existential motivation and emotional violence.

Participation rates across organisations

Tables A2.1 and A2.2 below show that participation and administration of the survey differed strongly between organisations. The questionnaire had not been administered at T4 by the majority of organisations. Two organisations hardly provided participants. For ORG.UK5, only one participant at T0 was entered, but no

(ex-)partner. For ORG.UK2, only one client and one ex-partner have answered the questionnaire at T0. Three organisations did not administer the questionnaire to clients at T0. ORG.UK3 neither administered the questionnaire at T2, nor included any (ex-)partner reports.

Table A2.1. Participation of male clients per organisation, across all timepoints.

	T0	T1	T2	T3	T4	Total
ORG.UK1	0	50	18	9	0	
ORG.UK2	1	0	0	0	0	
ORG.UK3	8	269	0	61	0	
ORG.UK4	44	0	20	13	0	
ORG.UK5	1	0	0	0	0	
ORG.IT1	20	2	17	19	5	
ORG.IT2	6	4	4	3	0	
ORG.CR1	0	6	6	6	0	
ORG.BU1	0	2	6	5	3	
Total	80	333	71	116	8	

Table A2.2. Participation of female (ex-)partners per organisation, across all timepoints.

	T0	T1	T2	T3	T4	Total
ORG.UK1	0	24	9	7	0	
ORG.UK2	1	0	0	0	0	
ORG.UK3	0	0	0	0	0	
ORG.UK4	13	2	6	9	0	
ORG.UK5	0	0	0	0	0	
ORG.IT1	14	0	9	1	2	
ORG.IT2	6	4	4	2	0	
ORG.CR1	0	6	6	6	0	
ORG.BU1	1	2	6	5	2	
Total	35	38	40	30	4	

Continuous participation

In analysing change over time, it is essential that clients participate at multiple timepoints. Table A2.3 indicates that few clients participated at all time-points. Both in regards of treatment course and sample size, a comparison between T1 and T3 was initially deemed reasonable. However, the majority of participants who form $n = 72$, stem from ORG.UK3. This was an organisation that did not administer T2 and was

the only organisation that did not include (ex-)partner-reports. For the research question, (ex-)partner-reports are essential. Deleting ORG.UK3, led to losing n = 53 and comparison between T1 and T3 could only be executed for a remaining n = 19.

Table A2.3. Original timepoints continuous participation.

	T0	T1	T2	T3	T4
T0					
T1	7				
T2	36	29			
T3	34	78	53		
T4	5	0	8	8	

It was decided to pool clients together for T0, T1 and T2 for baseline, who were then deemed ‘completers’ if they also participated at T3. This led to the following participation rates across organisations, shown in Table A2.4 for clients and in A2.5 for (ex)partners. Both participation and attrition rates among clients were largest in ORG.UK3.

Table A2.4. Participation of male drop-outs and completers per organisation.

	Clients			
	Drop-outs (n=300)		Completers (n=115)	
	n	%	n	%
ORG.UK1	42	14.0	9	7.8
ORG.UK2	1	0.3	0	0.0
ORG.UK3	218	72.7	59	51.3
ORG.UK4	32	10.7	15	13.0
ORG.UK5	1	0.3	0	0.0
ORG.IT1	2	0.7	18	15.7
ORG.IT2	3	1.0	3	2.6
ORG.CR1	0	0.0	6	5.2
ORG.BU1	1	0,3	5	4,3

Table A2.5. Participation of female drop-outs and completers organisation.

	(Ex-)Partners			
	Drop-outs (n=43)		Completers (n=21)	
	n	%	n	%
ORG.UK1	10	23.3	3	14.3
ORG.UK2	1	2.3	0	0.0
ORG.UK3	0	0.0	0	0.0

ORG.UK4	14	32.6	4	19.0
ORG.UK5	0	0.0	0	0.0
ORG.IT1	13	30.2	1	4.8
ORG.IT2	4	9.3	2	9.5
ORG.CR1	0	0.0	6	28.6
ORG.BU1	1	2.3	5	23.8

Results

Frequencies of biographical information are shown in Table A2.6. Chi-square tests indicated there were considerable differences between drop-outs and completers in age and referral route. There were significantly more young (18 to 30) and old (51 to over 60) drop-outs, $X^2(5) = 17.24, p = <.01$. There were also more mandatory referred completers and more voluntarily referred drop-outs, $X^2(2) = 11.23, p = <.01$. Differences in employment were marginally significant, with more completers who worked full-time and more drop-outs who were unemployed, $X^2(4) = 9.14, p = .06$.

Table A2.6. Demographic statistics (age, referral route compressed and detailed) of clients at baseline.

Variable	Drop-outs (n=300)		Completers (n=115)	
	n	%	n	%
Age				
18-21	18	6,0	0	0
22-30	108	36,0	25	21,7
31-40	96	32,0	40	34,8
41-50	53	17,7	27	23,5
51-60	16	5,3	12	10,4
over 60	7	2,3	3	2,6
Missing	2	,7	8	7
Employment				
Full-time	169	56,3	80	69,6
Part-time	16	5,3	8	7
Unemployed	58	19,3	11	9,6
Sickness leave/retirement	37	12,3	9	7,8
Something else	20	6,7	7	6,1
Income				
Struggling to pay for the essentials	42	14,0	19	16,5
Managing to pay for the essentials, no left over	89	29,7	30	26,1

Managing to buy the occasional treat or save sometimes	82	27,3	32	27.8
Managing regular treats and saving or holiday	24	8,0	18	15.7
Comfortably managing	54	18,0	15	13
High income	8	2,7	1	0.9
Referral route compressed (multiple responses possible)	1	,3		
Mandatory	40	13,3	27	23,5
Voluntary	231	77,0	64	55,7
Other	23	7,7	12	10,4
Missing	0	0	12	10,4

There were also significant differences in couple participation and motivation.

Significantly more completers (34.8%) than drop-outs (9.0%) participated with an (ex-)partner, $X^2(1) = 40.819, p = <.001$. Groups were similar in their relationship status and hopes, see Table A2.7. Completers further had significantly higher average existential motivation ($M = 3.30, SD = 2.44$) than drop-outs ($M = 2.70, SD = 2.59$), $t(409) = -2.15, p = .03$. Groups were also similar in regards of their self-reported use of emotional violence at baseline, as shown in Table A2.8.

Table A2.7. Relationship status and hopes for male drop-outs and completers.

	Drop-outs (n=300)		Completers (n=115)	
	n	%	n	%
Relationship status				
Together	159	53,0	58	50.4
Apart	107	35,7	47	40.9
Other	32	10,7	8	7
Missing	2	,7	2	1.7
Relationship hopes				
Together	169	56,3	65	56.5
Apart	23	7,7	13	11.3
Other	106	35,3	35	30.4
Missing	2	,7	2	1.7

Table A2.8. Emotional violence at baseline for male drop-outs and completers.

	Drop-outs	Completers	t	df	p
	(n=300)	(n=115)			
	M (.SD.)	M (.SD.)			
Existential motivation	2.87 (2.81)	3.60 (2.72)	-2.36	409	.02

Emotional violence	1.23 (.23)	1.24 (.29)	-.62	410	.54
--------------------	------------	------------	------	-----	-----

(Ex-)Partners were only asked about their relationship and IPV. Table A2.9 presents comparisons between female drop-outs and completers in regards of their relationship status and hopes. Table A2.10 contains mean scores on IPV for drop-outs and completers. Significance tests found them to be comparable in regards of their relationship status and hopes, as well as reported frequency of emotional violence by the client. However, the sample size was a lot smaller which makes significance testing less reliable.

Table A2.9. Relationship status and hopes for female drop-outs and completers.

	Drop-outs (n=42)		Completers (n=21)	
	n	%	n	%
Relationship status				
Together	25	59.5	13	61.9
Apart	11	26.2	7	33.3
Other	6	14.3	1	4.8
Missing	0	0	0	0
Relationship hopes				
Together	16	38.1	10	47.6
Apart	8	19.0	3	14.3
Other	17	97.6	8	38.1
Missing	1	2.4	0	0

Table A2.10. Emotional violence for female drop-outs and completers.

	Drop-outs (n=42)	Completers (n=21)
	M (S.D.)	M (S.D.)
Emotional violence	1.67 (.57)	1.58 (.50)

Discussion

Based on these results, drop-outs differed from completers in a few ways. They were often either in their twenties or over fifties compared to completers. Drop-outs were more often voluntarily referred and displayed a lower existential motivation than completers. These differences were however not reflected in differences their self-

reported baseline use of IPV. This does not imply the groups were also similar in their decrease of violence, as drop-outs' use of violence was not measured at follow-up. Therefore, it cannot be excluded that attrition is still predictive of recidivism (Jewell & Wormith, 2010) and the groups *did* differ in aspects that were unmeasured here. On top of that, the current study was not able to investigate characteristics of perpetrators who were not admitted or who were excluded before the start of the programme. Nor were the reasons for dropping out measured. Measuring these characteristics have been underlined as important aspects of attrition analysis (Lilley-Walker et al., 2016).

Appendix 3 - Factor analysis

Factor analysis of motivation to participate

To explore the variability and structure of items measuring motivation to participate in the programme, all 12 items were subjected to principal component factor analysis with varimax rotation, as seen in Table A3.1. Motivation was asked at T0, T1 and T2 and factor analysis was performed for the pooled baseline sample of those three timepoints, resulting in $n = 398$. The Kaiser-Meyer-Olkin measure confirmed sampling adequacy of $KMO = .87$, above the recommended threshold of $.60$. Bartlett's test of sphericity was significant, ($\chi^2(66) = 1534.56, p < .001$), indicating that at least two items correlate significantly and thus are suitable for dimension-reduction (Tobias & Carlson, 1969). The factor analysis initially yielded a three-factor solution, explaining 56.0% of the variance. The analysis did not support the theoretical distinction in functional and existential motivation within the sample.

The first dimension contains seven existential, personal/relational motives and one functional, relational motive, accounting for 36.6% of the variance, indicating strong reliability of the sub-scale. Cronbach's alpha is $\alpha = .88$, and after deleting the functional motive it is $\alpha = .87$. The second dimension contains functional, institutional-child items and accounts for 10.0% of the variance. But the correlation between the two items is negative, small and significant; $r = -.13, p < .01$. The correlation is negative because participants who are sent by the family court cannot be sent by the child protection services and vice versa. The third dimension entails functional, institutional-criminal motives, accounting for 9.5% of the explained variance. The correlation between the institutional-criminal items is positive, small and not significant, $r = .08$. As a result, the seven existential items are combined as a scale in a sum-variable and the remaining items are excluded.

Table A3.1. Results from factor analysis on motivation to attend programme.

Items	Factor			Dimension
	1	2	3	
I don't want my partner to be afraid of me	0.84	-0.09	0.04	Existential and relational
I want my partner/ex to feel safe around me	0.83	0.06	0.02	
I want to stop using abusive behaviour	0.78	0.02	-0.02	
I don't want my children be afraid of me	0.75	0.06	-0.17	
I want to stop using violence	0.73	0.10	0.08	
I want my relationship to be better	0.71	-0.17	-0.01	
I don't want my partner to leave me	0.64	-0.34	0.13	
I want to be a better parent to my children	0.55	0.15	-0.26	Institutional - Child
I have to come because the family court told me to	-0.03	0.70	-0.44	
I have to come because the child protection services told me to	-0.29	-0.49	0.04	Institutional - Criminal
I don't want to go back to prison again	0.13	0.20	0.65	
I have to come as part of my criminal court sentence or bail or parole conditions	0.03	0.48	0.58	

Notes. Extraction method; principal component analysis; No rotation. Factor loadings larger than 0.4 are in bold.