



# The European Journal of Psychology Applied to Legal Context

<https://journals.comadrid.org/ejpalc>



## Exploring Typology Categorizations of Male Perpetrators: A Methodology Study

Emily N. Weber<sup>a</sup>, Ashley R. Taylor<sup>a</sup>, Arthur L. Cantos<sup>b</sup>, Barbara G. Amado<sup>c</sup>, and K. Daniel O'Leary<sup>d</sup>

<sup>a</sup>Rosalind Franklin University of Medicine and Science, North Chicago, IL, USA; <sup>b</sup>University of Texas Rio Grande Valley, Edinburg, Texas, USA; <sup>c</sup>Universidad de Santiago de Compostela, Spain; <sup>d</sup>Stony Brook University, Stony Brook, NY, USA

### ARTICLE INFO

#### Article history:

Received 28 February 2019  
Accepted 5 May 2019

#### Keywords:

Intimate partner violence  
IPV (family only)  
Generally violent  
Self-report

### ABSTRACT

Intimate partner violence (IPV) perpetrators were categorized based on whether they were generally violent (GV) or family only violent (FO) using self-report or arrest records. Classification criteria to assess recidivism in perpetrators of IPV were evaluated herein to determine the incremental validity of using a perpetrator's criminal history in addition to their self-report information for categorization purposes. The concordance rates for categorizing subtypes of male perpetrators were compared for two methods, namely, self-report versus criminal history data. Categorizations were made based on self-reported history of violence and federal criminal records separately. Between measures consistency was defined as whether or not the self-report categorizations matched federal criminal record categorizations. It was hypothesized that self-report would not be sufficient as the sole method of categorizing male perpetrators, and the use of criminal history data would add to the validity of the categorization system. Self-reports of aggression were higher than criminal records of aggression. Using data sources together may yield the best outcomes for offenders and society. Implications are discussed.

### La exploración de las categorizaciones de la tipología de los varones violentos en la pareja: un estudio metodológico

### RESUMEN

Se clasificaron los varones que ejercen violencia en las relaciones de pareja (VP) en función de si eran violentos en general (VG) o solo en el entorno familiar (VF), empleando registros de autoinformes o de arrestos. Se analizaron los criterios de clasificación para evaluar la reincidencia de los infractores de VP con el fin de determinar la validez incremental del uso de los antecedentes penales del infractor, además de la información procedente de su autoinforme para la clasificación. Se compararon los índices de concordancia para categorizar los subtipos de infractores masculinos para dos métodos: los datos procedentes de autoinforme y los de antecedentes penales. La categorización se basó en la historia de violencia autoinformada y en los antecedentes penales por separado. La congruencia entre medidas se definió como la coincidencia o discrepancia de la categorización de autoinforme con la categorización de antecedentes penales. Se planteó la hipótesis de que el autoinforme no bastaba como único método para clasificar a los infractores masculinos y que el uso de datos procedentes de antecedentes penales aumentaba la validez del sistema de categorización. Hubo más autoinformes sobre agresión que antecedentes penales de agresión. El uso conjunto de ambos podría tener mejores resultados, tanto para los delincuentes como para la sociedad. Se discuten las implicaciones de estos resultados.

Data from a nationally representative survey suggests that 33% of women and 28% of men have experienced some form of physical violence by an intimate partner in their lifetime (Breiding, Chen, & Black, 2014). Research has focused on determining the characteristics and correlates of perpetrators in an attempt to capture the heterogeneity of perpetrators and subsequently reduce overall levels of violence (Bell & Naugle, 2008; Capaldi, Knoble, Shortt, & Kim, 2012). Understanding the processes involved in obtaining and documenting information regarding violence history and how this information is utilized is important to consider, as recognition of heterogeneity

within this population may improve the ability to predict treatment outcome (Stoops, Bennett, & Vincent, 2010). Many efforts have been made to identify and categorize different types of male perpetrators of intimate partner violence (IPV; e.g., Holtzworth-Munroe & Stuart, 1994). Other studies have attempted to identify batterer subtypes to increase the effectiveness of interventions with a variety of aggression subtypes (Boyle, O'Leary, Rosenbaum, & Hassett-Walker, 2008; Cunha & Gonçalves 2013; Holtzworth-Munroe & Stuart 1994; Langhinrichsen-Rohling, Huss, & Ramsey 2000). Research outcomes provide evidence for a relatively simple dichotomous categorization

Cite this article as: Weber, E. N., Taylor, A. R., Cantos, A. L., Amado, B. G., & O'Leary, K. D. (2019). Exploring typology categorizations of male perpetrators: A methodology study. *The European Journal of Psychology Applied to Legal Context*, 11, 93-97. <https://doi.org/10.5093/ejpalc2019a5>

Correspondence: [arthur.cantos@utrgv.edu](mailto:arthur.cantos@utrgv.edu) (A. L. Cantos).

ISSN: 1889-1861/© 2019 Colegio Oficial de Psicólogos de Madrid. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

of men as either generally violent (GV) or family only violent (FO), as a method that allows for more individually focused interventions based on characteristics of violence profiles (Cantos, Goldstein, Brenner, O'Leary, & Verborg, 2015; Goldstein, Cantos, Kosson, Brenner, & Verborg, 2015; Juarros-Basterretxea, Herrero, Fernández-Suárez, Perez, & Rodríguez Díaz, 2018).

Established methodologies used to categorize subtypes of perpetrators vary based on their theoretical or empirical grounds (Langhinrichsen-Rohling et al., 2000). The most widely utilized method across IPV research has been self-report, which has been used to categorize perpetrators based on certain dimensions (Stoops et al., 2010; Walsh et al., 2010; Waltz, Babcock, Jacobson, & Gottman, 2000). The self-report inventories used to make these categorizations include the Conflict Tactics Scale (Holtzworth-Munroe, Meehan, Herron, Rehman, & Stuart, 2000; Mauricio & Lopez 2009; Waltz et al., 2000), the Offender Assessment Tool (Stoops et al., 2010), psychopathology measures (Cunha & Gonçalves 2013; Holtzworth-Munroe et al., 2000; Langhinrichsen-Rohling et al., 2000; Walsh et al., 2010; Waltz et al., 2000), various personality assessment questionnaires (Langhinrichsen-Rohling et al., 2000; Walsh et al., 2010; Waltz et al., 2000), and other measures of aggression (Boyle et al., 2008; Cantos, Brenner, Goldstein, O'Leary, & Verborg, 2015). However, little research has focused on the validity of using such self-report information to categorize IPV male perpetrators (Heckert & Gondolf 2000). This is particularly important given the identification of a systematic source of error resulting from defensive responses in some contexts as custody litigants, personnel selection, and perpetrators (Arce, Fariña, Seijo, & Novo, 2015). Alternatively, some studies have used more objective measures, such as arrest records and police reports, to make these categorizations (Cantos & O'Leary, 2014; Stoops et al., 2010; Walsh et al., 2010). However, the incremental validity of using criminal history measures instead of, or in conjunction with, self-report has not been comprehensively assessed. The possibility of mis-categorizing men as "family only violent" when they are really "generally violent" has major implications in regards to the judgment of severity of risk and which intervention is most appropriate (Heckert & Gondolf, 2000). It is well established that psychological measurement of constructs includes some error that may bias estimates of reliability and true relationships (Schmidt, Le, & Ilies, 2003). This is potentially problematic for measures that rely on self-reported information to make accurate categorizations of GV or FO. Evidence suggests that male perpetrators tend to underreport or minimize violence in self-report measures (Browning & Dutton, 1986). Furthermore, agreement between partner reports of violence compared to offender reports tends to be moderate to low (Pearson correlation for physical aggression perpetrated by men is .43 and by women is .41; O'Leary & Williams, 2006). In consequence, the classification of perpetrators would rest on measures influenced by systematic sources of errors (Podsakoff, MacKenzie, Lee, & Podsakoff, 2003). The additional validity of using objective measurement of violence to make classifications is important to consider and can be compared by assessing categorization error for self-reported behavior and arrest records.

With respect to research that looks at criminal offenses overall (not just IPV), three sources are typically used to measure behavior: victimization surveys, self-report surveys, and official data from law enforcement (Kirk, 2006). Interestingly, self-report data has been found to produce higher estimates of criminal behavior and frequency of offending (Maxfield, Weiler, & Widom, 2000). Some have argued that self-report may actually be a more accurate method nearer to genuine criminal behavior as compared to other methods (Farrington, 2001; Thornberry & Krohn, 2003). However, prior research indicates that both self-report methods and official records methods have advantages and disadvantages. Self-report allows for a comprehensive collection of information on individual, familial, and environmental influences on criminal behavior, but may contain

recall error biases or response falsification (Kirk, 2006). Alternatively, arrest records contain specific and comprehensive information about criminal events. However, some argue they underestimate the true volume of crime due to underreporting from victims and perpetrators not being arrested (Kirk, 2006). It is also important to keep in mind that police discretion plays a role in who gets arrested, which arrests are recorded, and which charges are filed (Allen, 1984). These are a few of the many factors that contribute to measurement error when using police records (Maxfield et al., 2000) and the systematic sources of errors in behavioral research pointed out by Podsakoff et al. (2003).

Overall, research on general criminal populations found consistent results between self-report and official records when measuring offending behavior, with self-report yielding higher offense frequencies (Hindelang, Hirschi, & Weis, 1979; Kirk, 2006; Maxfield et al., 2000). Huizinga and Elliott (1986) reviewed the early studies on delinquent behavior and found that test-retest reliabilities for self-reported delinquent behavior were reported to range from .85 to .99. However, validity is much harder to assess given that there is no actual "gold standard" against which to judge (Thornberry & Krohn, 2000). In general, the consensus is that using multiple data sources is likely to be a more valid indicator of violence than results from a single source (Farrington, Loeber, Stouthamer-Loeber, Van Kammen, & Schmidt, 1996).

The aim of this study was to assess the validity of using two distinct measurements of violence to make perpetrator categorizations, by comparing 1) self-reported violent behavior alone and 2) National Law Enforcement Agencies Data System, which represent official arrest records (termed "LEADS") alone. Data originally gathered from a sample of perpetrators of intimate partner violence on probation in Lake County, IL was used in the current analyses (Cantos, Brenner et al., 2015). Categorizations of family only violent (FO) and generally violent (GV) in the previous study were made using a combination of sources. In this study, the likelihood of miscategorizing men was assessed by comparing self-report versus arrest records methods of categorization. Comparing the prevalence rates of each method will offer insight into which method is most useful in making accurate categorizations and for what purposes.

## Method

### Sample

The original sample consisted of 456 men on probation in Lake County, IL from 2006 to 2008. Our sample was a subset of the original sample with 385 men on probation during this same time. Men were between the ages of 17 and 72, with a mean age of 34.01 ( $SD = 10.78$ ). Thirty-four percent of the men reported themselves as single, 25.3% as having a girlfriend, 31.3% as married, and 8.6% as divorced. Fifty four percent of the men reported they were working and 45.2% unemployed. The majority of the men were Caucasian (45.7%), followed by 34.5% African American, 19.2% Latinos, and 0.5% Asian/Pacific islanders.

### Measures

The criteria used for categorizing men according to type of violence were based on a previously developed categorization system by Cantos, Goldstein et al. (2015). Other members of our research team had previously categorized the men in this sample as generally violent (GV) or family only violent (FO). The data used to make the initial categorizations were acquired from each participant's file, and included their Level of Service Inventory-Revised (LSI-R), Pre-Intake Probation Form, police reports, arrest records information, and psychological reports. For more detail on these measures, see earlier paper by Cantos, Brenner et al. (2015).

Men were categorized as FO if their file indicated no other record of past violent behavior. Men whose arrests consisted of traffic violations, drug offenses, and/or only domestic violence related offenses were also categorized as FO. Alternatively, men were categorized as GV if their file indicated a history of one or more aggressive acts against a non-intimate partner including battery, aggravated assault with or without a deadly weapon, armed robbery, disorderly conduct, or sexual assault. Resisting arrest was not sufficient to warrant a GV categorization if it was not in conjunction with an aforementioned arrest. Further, a history of aggression problems or gang affiliation in childhood, as indicated on the intake form, would be used to clarify categorizations of men as GV where categorizations made from criminal history data were unclear (i.e., battery arrest record without further qualification).

In the current study, the same criteria were utilized to categorize men. However, they were categorized twice separately using information from two different methods. The first method of categorization involved using only self-reported information from their intake assessment with a probation officer to make either a GV or FO categorization. This included information about previous acts of violence and gang membership. The second method used only arrest records to make the categorization. This contained information from all law enforcement agencies nationally and thus serves as a comprehensive summary of a perpetrator's criminal activity. Each of these methods (self-report vs. arrest records) was analyzed as a means of detecting FO or GV men.

## Coding

Two researchers were involved in categorizing perpetrators using both methods. The raters separately categorized each perpetrator by using self-report information only, and then made second independent categorizations using arrest records information only. Prior to coding, raters independently categorized a sample of the same 20 men for both sources of information to establish inter-rater reliability. The kappa for the preliminary cases was 1.0. Inter-observer drift was subsequently assessed by rating 20 men conjointly after the coding of 100 men. Raters continued to overlap on 20 subjects for every 100 subjects coded. No observer drift was noted and kappa's ranged from .875 to .894 ( $M = 0.885$ ).

In order to obtain a measure of coding reliability with respect to the use of the criteria to classify the perpetrators, true kappa ( $\bar{k}$ ) was calculated, given that the variables were categorical. True kappa is calculated like Cohen's original kappa which corrects for chance agreement, but is incomplete if the exact correspondence between the ratings is not verified and identifies the true concordance (Arce, Fariña, & Fraga, 2000). For example, if a perpetrator is classified as GV for an arrest A by rater 1 and also classified as GV for a different arrest by rater 2, the original kappa would classify this as concordance when in reality there is no exact correspondence in the coding and it would represent two episodes of non-concordance. Inter-rater reliability is usually obtained between raters, but this is also insufficient since reliability is not measured taking the passage of time (test-retest) into consideration nor between different raters. As a result, inter-rater reliability has to be measured in conjunction with intra-rater reliability (test-retest) and inter-context reliability (with other raters in other contexts), in order to estimate if different raters who are similarly trained in the coding system would obtain similar (concordant) results (Monteiro, Vázquez, Seijo, & Arce, 2018). In the present study, two raters coded all of the protocols (half each) and they each re-rated 10% of the original protocols after 10 days. The results reveal a very high true concordance of the inter-rater ratings,  $\bar{k} = .95$ , and inter-rater reliability,  $\bar{k} = .60$ . In addition, one of the raters was reliable in the ratings for a different study, inter-context reliability (Mach, Cantos, Weber, & Kosson, 2017). The results of the

true concordance are interpreted as high ( $> .61 \bar{k} < .81$ ), very high ( $\geq .81$ ), and weak ( $< .60$ ). Having established that the true reliability is very high in our study ( $\geq .81$ ), both intra-rater, as well as inter-rater and inter-context, we can conclude that the coding was completed reliably with the assigned criteria (Monteiro et al., 2018). Preliminary analysis consisted of the phi coefficient to evaluate the concordance rates between the categorizations made using only self-report information or arrest records. Subsequently, separate chi-square tests were used to compare the prevalence rates using the two methods.

## Results

### Between Measures Categorization Convergence

The convergence rate for categorizing men using self-report only versus categorizing men using arrest records only was significant,  $\phi = .392$ ,  $p < .001$ , but insufficient for the negative implications of a misclassification as 49.6% of self-reports' and arrest records' distributions are independent (i.e., non-overlapped;  $U1 = .496$ ). In consequence, only around 50% of perpetrators are classified in the same category by these methods. Classification rates made by arrest records and self-reports were contrasted by binomial tests.

### Rate Comparison for Self-Report vs. Arrest Records

In order to determine the extent of differences associated with making categorizations based on different methods, the cell counts from the four-fold table were used, and they are presented in Table 1. The data in Table 1 were first analyzed using chi square, exhibiting significant differences,  $\chi^2(1) = 59.10$ ,  $p < .001$ , and of a moderate magnitude,  $\phi = .39$ , in the classification of GV and FO perpetrators by arrest records and self-reports. Classification rates made by arrest records and self-reports were contrasted by binomial tests.

**Table 1.** Arrest Records Categorization by Self-Report Information Categorizations

Self-Report information	Arrest records		Total
	GV	FO	
GV	81	62	143
FO	45	197	242
Total	126	259	385

Note. GV = generally violent; FO = family only.

First, arrest records were considered as a gold standard of classifying FO or GV men. Arrest records classified 126 men as GV and 259 men as FO. However, as for GV men, self-reports failed to detect 45 of the 126 (i.e., 35.7%; these were detected by arrest records as GV, while self-reports failed classifying them as FO), a significant misclassification ( $p < .001$ ). In relation to those classified as FO by arrest records, 197 of the 259 men, i.e., 76.1%, were also classified as such by self-reports, a significant between-methods agreement ( $p < .001$ ).

Second, self-reports were considered as a gold standard of classifying FO or GV men. Self-reports classified 143 men as GV and 242 men as FO. However, as for GV men, arrest records failed to detect 62 of the 143, i.e., 43.4% (perpetrators informed – self-reported – about themselves as GV, while arrest records failed classifying them as FO), a significant misclassification ( $p < .001$ ). In relation to those classified as FO by self-reports, 197 of the 242 men, i.e., 81.4%, were also classified as such by self-reports, a significant between-methods agreement ( $p < .001$ ).

Comparatively, the misclassification rate is equal (if the 95% CIs for the observed proportion overlap, it indicates no mean

differences) for self-reports, .357, 95% CI [.27, .45], and arrest records, .434, 95% CI [.34, .52]. Likewise, the correct classification rate is equal for self-reports, .814, 95% CI [.74, .88], and arrest records, .761, 95% CI [.68, .84].

## Discussion

The aim of this study was to determine the extent to which one can categorize male perpetrators of intimate partner violence as generally violent (GV) or family only violent (FO) using methods of self-report versus official arrest records, and whether or not self-report alone is sufficient to make these categorizations. Our results indicate a lack of consistency across both methods of categorization. There was a modest correlation, significant but insufficient, between the categorizations made using only self-report compared to the arrest records. Around half of the perpetrators are classified in the same category by these methods and half of the perpetrators would be misclassified by either method. Possible explanations for this inconsistency across methods could be explained by men potentially under-reporting previous acts of violence. This would result in an FO categorization when, in fact, their violence profile based on official arrest records reflects a GV categorization. Previous research has documented the tendency for men to under-report their history of violence compared to the reports of violence from their partners (Browning & Dutton 1986; O'Leary & Williams, 2006), which supports this notion of miscategorization. Alternatively, arrest records may not fully capture a perpetrator's past violent behavior, including things such as gang membership, physical fights in school, or other violent behavior that was never officially charged. Looking at these methods in isolation would create conflicting category profiles. These discrepancies lend support towards using both methods together to accurately categorize male perpetrators. However, it is important to consider the remaining error inherent with using both sources of information. Using both arrest records and self-report information does not eliminate all sources of error, as some perpetrators may not self-report generalized aggression or may never have been arrested for these aggressive crimes.

In evaluating categorization error, although rates of miscategorization and correct classification rates are similar for both methods, our results indicated that GV categorizations are lower when using arrest records compared to self-report methods. In short, some men self-reported more interpersonal violence than conveyed in their official records. FO miscategorizations, men categorized as FO when they are GV, downplay pervasive violence history and violence potential. Additionally, men who are categorized as GV tend to have lower rates of treatment completion (Cantos, Goldstein et al., 2015; Fowler & Western, 2011; Huss & Ralston, 2011; Langhinrichsen-Rohling et al., 2000; Rooney & Hanson, 2001) as well as higher rates of post-probation recidivism (Cantos, Brenner et al., 2015; Cantos, Kosson, Goldstein, & O'Leary, in press). Failing to accurately categorize GV men as GV may result in an under-calculation of the level of risk this offender poses to society. Finally, perpetrators may not be accurately categorized as GV when using self-report information alone, as some men do not self-report violence, but have been arrested for aggression towards others. This miscategorization is supported by the low convergence rates between both methods and the chi-square analyses using arrest records as the criterion to categorize men as GV.

Based on these results, it is recommended to use self-report in conjunction with official arrest records in order to best categorize male perpetrators of intimate partner violence. However, if only one method is available, relying on self-report would most likely minimize the risk to victims. While self-report might be the most practical and convenient method to determine categorizations, reliance on self-report information alone may be doing perpetrators, and their potential victims, a disservice, and could result in

categorization errors based upon these findings. If men are going to be referred to specific treatment programs based upon these categories, it is imperative to make certain that they are categorized correctly and that efforts are being made to minimize error. Accurate categorization is an integral part of establishing focused and effective intervention strategies relevant to type-specific characteristics. Understanding what information is valid and reliable to use in the categorization process gives us insight into how to best categorize perpetrators of IPV, and how to best guide treatment. If categories are assigned without focusing on the validity of these categorical methods, then treatment outcomes and program design may not be accurately targeting the specified group of men they are intended to benefit. The results yield important implications for the use of self-report measures as the "gold standard" of categorization. Historically, very few studies utilize additional objective measures, such as police reports and arrest records (Stoops et al., 2010; Walsh et al., 2010). These results indicate that the most accurate and ideal means when aiming to make correct categorizations of men should include a combination of both self-report and objective measurements of a man's aggressive behavior.

Despite the wealth of knowledge offered by these results, they must be taken into consideration with caution given the limitations of the study. The sample was a sample of male probationers, and thus the results may not necessarily extend to community samples or other dissimilar samples. Furthermore, not all male perpetrator populations have information available from national crime databases, which was unique to this sample. Finally, as noted above, using both sources of information does not guarantee that all error risk has been eliminated from miscategorizing perpetrators. Overall, this study is the first to our knowledge to evaluate methodology and present evidence for validity considerations in the context of categorizing intimate partner violence perpetrators, specifically as it pertains to self-report information. Detection of incremental validity in making accurate categorizations was assessed using the cross-validation of different methods of constructs. This was important to assess given the broad acceptance of using only self-report methods to categorize male IPV perpetrators. Categorizing male IPV perpetrators dictates the consequences for these men, which underscores the importance of making accurate distinctions for the perpetrator and for society.

## Conflict of Interest

The authors of this article declare no conflict of interest.

## Acknowledgements

We would like to thank all those who assisted in this project. We would like to give a special thanks to the staff at Lake County Department of Probation for their support and assistance of our lab's research involving intimate partner violence. Thanks to Gabriela Ontiveros for her help in revising the manuscript.

## References

- Allen, R. J. (1984). Impact of arrest records on the exercise of police discretion. *Law and Contemporary Problems*, 47, 287-302. <https://doi.org/10.2307/1191694>
- Arce, R., Fariña, F., & Fraga, A. I. (2000). Género y formación de juicios en un caso de violación [Gender and juror judgment making in a case of rape]. *Psicothema*, 12, 623-628. Retrieved from <http://www.psicothema.com/pdf/381.pdf>
- Arce, R., Fariña, F., Seijo, D., & Novo, M. (2015). Assessing impression management with the MMPI-2 in child custody litigation. *Assessment*, 22, 769-777. <https://doi.org/10.1177/1073191114558111>
- Bell, K. M., & Naugle, A. E. (2008). Intimate partner violence theoretical considerations: Moving towards a contextual framework. *Clinical Psychology Review*, 28, 1096-1107. <https://doi.org/10.1016/j.cpr.2008.03.003>

- Boyle, D. J., O'Leary, K. D., Rosenbaum, A., & Hassett-Walker, C. (2008). Differentiating between generally and partner-only violent subgroups: Lifetime antisocial behavior, family of origin violence, and impulsivity. *Journal of Family Violence, 23*, 47-55. <https://doi.org/10.1007/s10896-007-9133-8>
- Breiding, M. J., Chen J., & Black, M. C. (2014). *Intimate partner violence in the United States—2010*. Atlanta, GA: National Center for Injury Prevention and Control, Centers for Disease Control and Prevention.
- Browning, J., & Dutton, D. (1986). Assessment of wife assault with the Conflict Tactics Scale: Using couple data to quantify the differential reporting effect. *Journal of Marriage and the Family, 48*, 375-379. <https://doi.org/10.2307/352404>
- Cantos, A., Brenner, L., Goldstein, D. A., O'Leary, K. D., & Verborg, R. (2015, November). *Results of a 5-year outcome study with perpetrators of intimate partner violence placed on probation and mandated to a treatment program in Lake County Illinois*. VIII Congreso Internacional y XIII Nacional de Psicología Clínica. Granada, Spain.
- Cantos, A. L., Goldstein, D. A., Brenner, L. H., O'Leary, K. D., & Verborg, R. J. (2015). Correlates and program completion of family only and generally violent perpetrators of intimate partner violence. *Behavioral Psychology, 23*, 549-569.
- Cantos, A. L., Kosson, D. S., Goldstein, D. A., & O'Leary, K. D. (in press). Treatment impact on recidivism of family only v generally violent partner violence perpetrators. *International Journal of Clinical and Health Psychology*.
- Cantos, A. L., & O'Leary, K. D. (2014). One size does not fit all in treatment of intimate partner violence. *Partner Abuse, 5*, 204-236. <https://doi.org/10.1891/1946-6560.5.2.204>
- Capaldi, D. M., Knoble, N. B., Shortt, J. W., & Kim, H. K. (2012). A systematic review of risk factors for intimate partner violence. *Partner Abuse, 3*, 231-280. <https://doi.org/10.1891/1946-6560.3.2.231>
- Cunha, O., & Gonçalves, R. A. (2013). Intimate partner violence offenders: Generating a data-based typology of batterers and implications for treatment. *European Journal of Psychology Applied to Legal Context, 5*, 131-139. <https://doi.org/10.5093/ejpalc2013a2>
- Farrington, D. P. (2001). *What has been learned from self-reports about criminal careers and the causes of offending?* (Report for the Home Office). London, UK: Home Office.
- Farrington, D. P., Loeber, R., Stouthamer-Loeber, M., Van Kammen, W. B., & Schmidt, L. (1996). Self-reported delinquency and a combined delinquency seriousness scale based on boys, mothers, and teachers: Concurrent and predictive validity for African-Americans and Caucasians. *Criminology, 34*, 493-517. <https://doi.org/10.1111/j.1745-9125.1996.tb01217.x>
- Fowler, K. A., & Western, D. (2011). Subtyping male perpetrators of intimate partner violence. *Journal of Interpersonal Violence, 26*, 607-639. <https://doi.org/10.1177/0886260510365853>
- Goldstein, D. A., Cantos, A. L., Kosson, D. S., Brenner, L. H., & Verborg, R. J. (2015). Perpetrator type moderates the relationship between severity of intimate partner violence and recidivism. *Criminal Justice and Behavior, 43*, 879-898. <https://doi.org/10.1177/0093854815616841>
- Heckert, D., & Gondolf, E. W. (2000). Assessing assault self-reports by batterer program participants and their partners. *Journal of Family Violence, 15*, 181-197. <https://doi.org/10.1023/A:1007594928605>
- Hindelang, M. J., Hirschi, T., & Weis, J. G. (1979). Correlates of delinquency: The illusion of discrepancy between self-report and official measures. *American Sociological Review, 44*, 995-1014. <https://doi.org/10.2307/2094722>
- Holtzworth-Munroe, A., Meehan, J. C., Herron, K., Rehman, U., & Stuart, G. L. (2000). Testing the Holtzworth-Munroe and Stuart (1994) batterer typology. *Journal of Consulting and Clinical Psychology, 68*, 1000-1019. <https://doi.org/10.1037/0022-006X.68.6.1000>
- Holtzworth-Munroe, A., & Stuart, G. L. (1994). Typologies of male batterers: Three subtypes and the differences among them. *Psychological Bulletin, 116*, 476-497. <https://doi.org/10.1037/0033-2909.116.3.476>
- Huizinga, D., & Elliott, D. S. (1986). Reassessing the reliability and validity of self-report delinquency measures. *Journal of Qualitative Criminology, 2*, 293-327. <https://doi.org/10.1007/BF01064258>
- Huss, M. T., & Ralston, A. (2011). Do batterer subtypes really matter? Treatment completion, treatment response, and recidivism across a batterer typology. *Criminal Justice and Behavior, 35*, 710-724. <https://doi.org/10.1177/0093854808316218>
- Juarros-Basterretxea, J., Herrero, J., Fernández-Suárez, A., Perez, B., & Rodríguez Díaz, F. J. (2018). Are generalist batterers different from generally extra-family violent men? A study among imprisoned male violent offenders. *European Journal of Psychology Applied to Legal Context, 10*, 8-14. <https://doi.org/10.5093/ejpalc2018v10n1a1>
- Kirk, D. S. (2006). Examining the divergence across self-report and official data sources on inferences about the adolescent life-course of crime. *Journal of Quantitative Criminology, 22*, 107-129. <https://doi.org/10.1007/s10940-006-9004-0>
- Langhinrichsen-Rohling, J., Huss, M. T., & Ramsey, S. (2000). The clinical utility of batterer typologies. *Journal of Family Violence, 15*, 37-53. <https://doi.org/10.1023/A:1007597319826>
- Mach, J. L., Cantos, A. L., Weber, E. N., & Kosson, D. S. (2017). The impact of perpetrator characteristics on the completion of a partner abuse intervention program. *Journal of Interpersonal Violence*. Advance online publication. <https://doi.org/10.1177/0886260517719904>
- Mauricio, A. M., & Lopez, F. G. (2009). A latent classification of male batterers. *Violence and Victims, 24*, 419-438. <https://doi.org/10.1891/0886-6708.24.4.419>
- Maxfield, M. G., Weiler, B. L., & Widom, C. S. (2000). Comparing self-reports and official records of arrests. *Journal of Quantitative Criminology, 16*, 87-110. <https://doi.org/10.1023/A:1007577512038>
- Monteiro, A., Vázquez, M. J., Seijo, D., & Arce, R. (2018). ¿Son los criterios de realidad válidos para clasificar y discernir entre memorias de hechos auto-experimentados y de eventos vistos en vídeo? [Are the reality criteria valid to classify and to discriminate between memories of self-experienced events and memories of video-observed events?]. *Revista Iberoamericana de Psicología y Salud, 9*, 149-160. <https://doi.org/10.23923/j.rips.2018.02.020>
- O'Leary, K. D., & Williams, M. C. (2006). Agreement about acts of aggression in marriage. *Journal of Family Psychology, 20*, 656-662. <https://doi.org/10.1037/0893-3200.20.4.656>
- Podsakoff, P. M., MacKenzie, S. B., Lee, J.-Y., & Podsakoff, N. P. (2003). Common method bias in behavioral research: A critical review of the literature and recommended remedies. *Journal of Applied Psychology, 88*, 879-903. <https://doi.org/10.1037/0021-9010.88.5.879>
- Rooney, J., & Hanson, R. K. (2001). Predicting attrition from treatment programs for abusive men. *Journal of Family Violence, 16*, 131-149. <https://doi.org/10.1023/A:1011106902465>
- Schmidt, F. L., Le, H., & Ilies, R. (2003). Beyond alpha: An empirical examination of the effects of different sources for measurement error on reliability estimates for measures of individual differences constructs. *Psychological Methods, 8*, 206-224. <https://doi.org/10.1037/1082-989X.8.2.206>
- Stoops, C., Bennett, L., & Vincent, N. S. (2010). Development and predictive ability of a behavior based typology of men who batter. *Journal of Family Violence, 25*, 325-335. <https://doi.org/10.1007/s10896-009-9294-8>
- Thornberry, T. P., & Krohn, M. D. (2000). The self-report method for measuring delinquency and crime. *Criminal Justice, 4*, 33-83. Retrieved from [https://www.ncjrs.gov/criminal\\_justice2000/vol\\_4/04b.pdf](https://www.ncjrs.gov/criminal_justice2000/vol_4/04b.pdf)
- Thornberry, T. P., & Krohn, M. D. (2003). Comparison of self-report and official data for measuring crime. In J. P. Pepper & C. V. Petrie (Eds.), *Measurement problems in criminal justice research: Workshop summary* (pp. 43-94). Washington, DC: National Academy of Sciences.
- Walsh, Z., Swogger, M. T., O'Connor, B. P., Chatav Schonbrun, Y., Shea, M., & Stuart, G. L. (2010). Subtypes of partner violence perpetrators among male and female psychiatric patients. *Journal of Abnormal Psychology, 119*, 563-574. <https://doi.org/10.1037/a0019858>
- Waltz, J., Babcock, J. C., Jacobson, N. S., & Gottman, J. M. (2000). Testing a typology of batterers. *Journal of Consulting and Clinical Psychology, 68*, 658-669. <https://doi.org/10.1037/0022-006X.68.4.658>

