

research

Determining Batterer Intervention Program Treatment Intensities: An Illustration Using the Ontario Domestic Assault Risk Assessment

Dana L. Radatz, PhD 

Niagara University, Lewiston, NY

N. Zoe Hilton, PhD, C Psych 

University of Toronto and Waypoint Research Institute, Penetanguishene, ON

To incorporate evidence-based practices into batterer intervention programs (BIPs), some intimate partner violence (IPV) researchers have begun to explore the integration of the principles of effective intervention (PEI) into BIPs. The PEI risk principle states that programs should assess offenders' risk for recidivism, and then match offenders' risk levels to the intensity of treatment and other interventions. Therefore, this study demonstrates how BIPs can use the Ontario Domestic Assault Risk Assessment (ODARA) to guide treatment intensity decisions for IPV offenders. Using data from men with criminal charges for IPV incidents in the original ODARA construction and validation research ($N = 970$), we illustrate how low, medium, and high categories can be created and used to inform BIP treatment intensity decisions (as outlined by the risk principle). Results indicate the ODARA can be used to inform a three-tiered categorical system for BIPs, in accordance with the risk principle of the PEI. Study results provide an avenue for BIP service providers to consider when determining treatment intensity for its IPV offenders. Suggestions for future research are also provided.

KEYWORDS: intimate partner violence; IPV; principles of effective intervention; ODARA; risk assessment; risk principle

INTRODUCTION

Many scholars and practitioners have proposed that batterer intervention programs (BIPs) should work to incorporate evidence-based practices within their programming as a viable avenue for increased effectiveness at reducing recidivism (Babcock

et al., 2016; Cannon, Hamel, Buttell, & Ferreira, 2016; Radatz & Wright, 2016; Stewart, Flight, & Slavin-Stewart, 2013). In this vein, one suggestion has been to explore evidence-based practices used in offender treatments within corrections, known as the risk-need-responsivity principles of effective intervention (PEI; Bonta & Andrews, 2016), into BIPs (Radatz & Wright, 2016; Stewart et al., 2013). The PEI serve as a guiding, evidence-based framework of standards for offender treatment, that when followed, have been shown to yield greater reductions in recidivism (e.g., Andrews, Bonta, & Hoge, 1990; Andrews, Zinger et al., 1990; Bonta & Andrews, 2016). Among the first of many principles within the PEI, the risk principle states that offenders must be assessed for their probability to reoffend, assigned a risk level, and then provided treatment with a level of intensity that corresponds, such as by classifying offenders into low, medium, and high treatment intensity categories (Bonta & Andrews, 2016). That is, relatively high-risk offenders are placed in a high priority, highly intensive correctional treatment program, and the lowest risk offenders are candidates for the least intensive treatment, less secure custody, lower levels of supervision, or minimal correctional intervention. While the risk principle derived from research studies on general offenders, researchers have posited that it could be applied to IPV offenders specifically (Radatz & Wright, 2016; Stewart et al., 2013).

IPV offender risk assessment has been extensively examined (e.g., Hilton & Eke, 2017); however, there is a dearth of research regarding the use of IPV risk assessment instruments to aid in the determination of offenders' treatment intensity assignments within BIPs (as outlined by the risk principle). In particular, it remains unclear how IPV risk assessment tools' scoring guidelines can be comported and used to inform BIP treatment intensity decisions. Therefore, the present study illustrates how an existing risk assessment tool, namely the Ontario Domestic Assault Risk Assessment (ODARA; Hilton et al., 2004; Hilton, Harris, & Rice, 2010), can be used by BIP service providers to facilitate the application of the risk principle to decisions about BIP treatment intensity for IPV offenders.

THE RISK PRINCIPLE AND INTIMATE PARTNER VIOLENCE TREATMENT

To our knowledge, only a few studies have examined the risk principle in relation to IPV offender treatment (Connors, Mills, & Gray, 2012, 2013; Gover, Richards, & Tomsich, 2015; Hansen, Davis, Smith, & Hilkey, 2016; Stewart, Gabora, Kropp, & Lee, 2014; Stewart & Power, 2014). Each of those studies employed the use, in part, of an established IPV risk assessment tool. Connors et al. (2012) and Connors et al. (2013) conducted separate evaluations of IPV treatment programs for Canadian incarcerated offenders who were screened into high-intensity and medium-intensity programs, using the well-established IPV assessment tool, Spousal Assault Risk Assessment (SARA; Kropp & Hart, 1997; Kropp, Hart, Webster, & Eaves, 1998, 1999). The high-intensity program had more modules covering additional content, more

group and individual sessions, and longer group sessions, than the medium-intensity program. Both programs were associated with desirable and significant pre–post changes in measures of attitudes and program-related skills and participation.

Other Canadian researchers, Stewart et al. (2014), conducted an outcome evaluation of similar moderate- and high-intensity treatment programs delivered in accordance with the PEI framework. The study revealed that IPV offenders who were assessed and subsequently matched with a corresponding treatment intensity program (as outlined by the risk principle) were less likely to recidivate compared to untreated IPV offenders. Specifically, offenders who attended the high-intensity treatment program had a significantly reduced likelihood to reoffend, and IPV offenders who attended the moderate-intensity treatment program also showed some improvement, albeit not significantly (Stewart et al., 2014).

In 2010, the State of Colorado revised its standards regarding court ordered IPV offender treatment to better align with evidence-based practices, particularly the risk, need, and responsivity principles of the PEI (Hansen et al., 2016). Under the new standards and true to the adherence of the risk and need principles, IPV offenders are to be administered a 14-item risk assessment, the Domestic Violence Risk and Needs Assessment (DVRNA; Hansen et al., 2016) to assess their risk for reoffending. Once the assessment instrument is scored, offenders are then classified and matched with the appropriate treatment intensity program (Level A—low intensity, Level B—medium intensity, Level C—high intensity; Gover, 2011). The DVRNA is derived from several widely recognized and empirically supported offender risk assessments, including the SARA, the ODARA (Hilton et al., 2004, 2010), the Domestic Violence Screening Instrument (DVSI; Williams & Houghton, 2004), the Danger Assessment (Campbell, Webster, & Glass, 2009), and the Level of Service Inventory (LSI, LSI-R; Andrews & Bonta, 1995, 2001). Hansen et al. (2016) conducted a process evaluation of the implementation of the revised Colorado standards and the use of the DVRNA. Their study indicated that the DVRNA did place IPV offenders in their appropriate categories (i.e., Level A, Level B, and Level C). Though the initial findings from these studies are promising, the relatively new approach of applying the risk principle to treatment intensity categories has yet to be widely applied to BIPs.

Unfortunately, as Radatz and Wright (2016) explain, though BIPs are conducting assessments, many do not use the information to classify and place offenders into corresponding treatment intensity programs, and instead opt to treat all IPV offenders as a homogenous group. Additionally, BIPs operating in states that do not have a unified set of standards for evidence-based practices (like those in Colorado) may encounter practical barriers (e.g., limited funds, time, and resources) for employing a trained mental health professional to conduct assessments on their IPV offenders. Therefore, in an effort to move toward the integration of the risk principle into BIPs, some BIPs may need an IPV risk assessment instrument that allows them to assess and subsequently categorize IPV offenders into treatment intensity categories without the use of mental health professionals.

INTIMATE PARTNER VIOLENCE RISK ASSESSMENT

An IPV risk assessment instrument that can be quickly and reliably scored, using the limited information available at treatment referral and intake, and without extensive professional training required, would be the most pragmatic for BIPs. Two structured and validated tools that would meet these criteria are the Brief Spousal Assault Form for the Evaluation of Risk (B-SAFER; Kropp, 2008) and the ODARA (Hilton et al., 2004, 2010). The B-SAFER is a shortened version of the SARA, with 10–15 items intended to be scored in police investigations. The B-SAFER score was significantly associated with IPV and nonviolent offenses against an intimate partner in a follow-up study of Spanish prisoners 15 months after release (Loinaz, 2014). In other studies, researchers found that the B-SAFER was not associated with IPV recidivism by men assessed by police officers in Sweden (Belfrage & Strand, 2012) or in court-mandated treatment in the United States (Gerbrandij, Rosenfeld, Nijdam-Jones, & Galietta, 2018). Belfrage and Strand attributed their results to the police taking more action in cases they assessed to be high risk. Further research revealed that police officers' risk management recommendations were related to lower recidivism among offenders assessed as high risk, but increased recidivism among those assessed to be low risk (Storey, Kropp, Hart, Belfrage, & Strand, 2014), which is consistent with evidence regarding the PEI.

Like the B-SAFER, the ODARA uses information gathered from policing and criminal justice agencies to determine an offender's risk for IPV recidivism. The ODARA was developed through empirical research including a follow-up of 589 men following an "index" assault on a female marital or cohabiting partner recorded in a police report (Hilton, Harris, & Rice, 2001). The ODARA consists of 13 items that were unique predictors of domestic violence recidivism in the research: preindex domestic assault, preindex non-domestic assault, preindex custodial sentence, failure on preindex conditional release, index threat to harm, index confinement of the partner, violence against others, substance use, more than one child altogether, partner has child from previous partner, partner concern about future assault, assault on partner while pregnant, partner barriers to support (Hilton et al., 2010). The total ODARA score is the sum of the items present, so that the ODARA has a possible range of 0–13.

In cross-validations of the ODARA's predictive accuracy for domestic violence recidivism in samples overlapping with the present study sample, area under the curves (AUCs) ranged from .65 to .74, considered medium to large effects (Rice & Harris, 2005). In a meta-analysis, the ODARA was reported to be the most accurate domestic assault risk assessment tool (average AUC = .67; Messing & Thaller, 2013). Subsequent research demonstrated the ODARA's ability to accurately estimate risk of domestic assault recidivism and other criminal outcomes by men who committed violence, sexual assault, or threats against their marital, cohabiting or dating partners (Gerth, Rossegger, Singh, & Endrass, 2015; Hilton & Eke, 2016; Jung & Buro, 2017; Olver & Jung, 2017; Rettenberger & Eher, 2013). Inter-rater reliability for the ODARA score has yielded correlation coefficients of at least .90 in original data used

for the present study (Hilton et al., 2004) and in other research (e.g., Jung & Buro, 2017; Olver & Jung, 2017).

Assessors can learn to score the ODARA accurately with minimal training (Hilton & Ham, 2015) and brief training is available free online (odara.waypointcentre.ca) making it an attractive assessment tool for BIPs to explore. In addition, the ODARA is an actuarial instrument, meaning that it provides a statistical estimate of offenders' risk, including their percentile rank, that shows how one individual compares with others in the known population with respect to risk. Actuarial data lends itself to classifying offenders into low-, medium-, and high-intensity treatment categories using percentiles rather than simply raw score or clinical judgment. Using percentiles means that treatment intensity cutoffs can be determined for a given population based on the availability of services and other resources. That is, if an agency knows that it can only provide its most intensive treatment program to half of the referred offenders, an actuarial risk assessment can be used to identify the offenders who score at or above the 50th percentile for their referral list.

Currently, the ODARA is the only actuarial assessment instrument that allows minimally trained individuals to identify risk for reoffending through the use of criminal justice agency information (e.g., police and probation reports) alone. Thus, the ODARA is a suitable choice for BIPs to explore as a tool for classifying offenders into treatment intensity categories at intake, and in the present study we illustrate how this can be done.

PRESENT STUDY

Given the fairly recent call for BIPs to move toward the implementation of evidence-based practices (Babcock et al., 2016; Cannon et al., 2016; Radatz & Wright, 2016; Stewart et al., 2013), there is limited research that explores applying the PEI to IPV offender treatment. In particular, little is known about how IPV risk assessment can be used to determine treatment intensity decisions for BIPs as outlined by the risk principle of the PEI. Therefore, we examined how the ODARA may be used to inform treatment intensity decisions within BIPs, by reanalyzing the original data used to develop the ODARA. Specifically, we aimed to illustrate how ODARA scores can be divided into low, medium, and high categories (in reference to an offender's risk for recidivism) and then subsequently utilized by BIP service providers as a starting point in determining treatment intensity categories for IPV offenders.

METHOD

We reanalyzed data that were used to develop and validate the ODARA, combining samples from three previous studies (Hilton et al., 2004; Hilton, Harris, Rice, Houghton, & Eke, 2008; Hilton & Harris, 2009a). Data collection was entirely archival, so there was no participant recruitment or informed consent procedure.

Police records management databases were used to identify cases from police occurrence reports that met the definition of an “index assault”: a man physically assaulted his current or former female domestic partner (married or cohabiting), or made a credible threat of death with a weapon in hand in her presence, as documented in a police occurrence report. Original data included 1,421 cases. For the present study we selected only those men who had been charged for a criminal offense at the time of the index assault, to most closely parallel the population of offenders who are diverted or sentenced to BIPs following a criminal charge for IPV. Main effects reported below were similar when examined in the entire sample including charged and noncharged offenders.

Sample

Of the 1,421 available cases, 970 of male offenders were charged with a criminal offense at the time of the index assault. These 970 men constitute the present sample. All cases had been identified through one of three police records management systems. One system was used across the Province of Ontario by approximately 50 municipalities as well as rural and First Nations policing services, and two systems were used by policing services in metropolitan regions in the Greater Toronto Area. Inclusion criteria were that each case involved a male accused of physically assaulting (or threatening death with a weapon in hand in the victim’s presence) a female victim with whom he was or had been married or cohabiting. One index assault was identified for each case, being the most recent domestic assault meeting the inclusion criteria up to December 31, 1996 (provincial system) or 1997 (metropolitan systems).

Variables

ODARA. We used the ODARA score available in the dataset. The data were originally used to create and validate the optimal prediction model, through a series of regression analyses of a large number of variables, to identify the strongest and most consistent predictors of domestic violence recidivism (Hilton et al., 2004).

The original dataset contained ODARA item scores and we calculated the raw ODARA total without prorating for missing information. ODARA risk categories were derived according to the published interpretation procedure (Hilton et al., 2010) such that each score from 0 to 4 represents a separate category of increasing risk, scores 5 and 6 comprise the second highest risk category, and scores 7 to 13 comprise the highest risk category.

Domestic Violence Recidivism. As in the original ODARA dataset, recidivism was defined as any postindex physical assault against a female domestic partner or former partner, recorded in a police occurrence report or criminal record up to 2004. This definition did not require criminal charges for the recidivism offense.

Procedure

Original data collection occurred at one time. This procedure allowed for a retrospective follow-up design with a follow-up period of approximately 5 years on average.

A primary coder with a graduate research analyst diploma was employed for the entire data collection period, and supervised by two experienced research psychometrists who coded part-time. All information was gathered from the police records management system and from a record of criminal charges and convictions maintained in a national police database. Information was quantified using a coding form with detailed instructions for determining whether the information met the criteria for each variable and value level.

Data for both the ODARA items and postindex domestic violence recidivism were collected and coded at the same time, without masking, due to the manner in which current and historic case information was interspersed within each occurrence report and criminal record. Masked inter-rater reliability was measured at the original data collection stage. Thirty randomly selected cases were masked by printing and cutting out or covering contaminating information, then coded independently by two coders, and only variables with inter-rater reliability coefficients of $r \geq .80$ or $\kappa \geq .70$ were used in the analyses to develop the ODARA (Hilton et al., 2004).

Data Analysis Plan

We conducted receiver operating curve analyses for the ODARA and postindex domestic violence recidivism for the 970 men with criminal charges at the index domestic assault, and examined the resulting AUC value to confirm that the ODARA was a statistically significant predictor of recidivism in this combined sample. We anticipated a medium effect size. We then examined the distribution of ODARA scores and divided them into “low,” “medium,” and “high” treatment intensity categories using two rules. First, we chose to select categories based on percentile distributions rather than rates of recidivism, because we wished to illustrate how BIPs can identify treatment intensity groups based on their own referral population without the need for population-specific recidivism statistics. Basing categories on percentiles required that we should not divide ODARA scores that are associated with the same percentiles (i.e., scores of 5–6 would not be split between two treatment intensity categories, nor would scores of 7–13); and the three categories should have approximately equal proportions of cases as indicated by percentiles (i.e., terciles or one-third of the sample each).

RESULTS

The 970 male offenders had a mean age of 36.9 years ($SD = 10.4$) at the index date, and their female partners $M = 33.7$ ($SD = 10.0$). Most were married ($N = 440$, 45%) or cohabiting ($N = 353$, 38%) with the victim at the index date. All offenders had incurred criminal charges related to the index assault, most with common assault ($N = 660$, 68%). The mean ODARA score was 3.20 ($SD = 2.08$, range 0–11). Within an average follow-up of 5.50 years after the index date ($SD = 1.54$), 333 (34%) committed a domestic violence recidivism offense.

TABLE 1. ODARA Scores and Domestic Violence Recidivism in a 5.5 Year Follow-Up of 970 Men With Criminal Charges at an Index Assault Against a Female Domestic Partner

ODARA score	<i>N</i>	Percent	Percent recidivism	Treatment intensity category
0	64	7	8	Low
1	142	15	19	Low
2	193	20	25	Low
3	207	21	35	Medium
4	138	14	37	Medium
5–6	151	16	50	High
7–13	75	8	72	High

Note. ODARA = Ontario Domestic Assault Risk Assessment.

ODARA score predicted domestic violence recidivism with a medium effect size, $AUC = .689$ ($SE = .018$), 95% CI = [.654, .724]. The distribution of ODARA scores and associated recidivism rates are shown in Table 1.

ODARA scores are raw scores without prorating for missing items (items with unclear or incomplete information were scored as 0). Domestic violence recidivism is defined as a postindex physical assault against a female domestic partner appearing in a police occurrence report or criminal record. Distribution and recidivism statistics may not generalize to different sample inclusion criteria, different ODARA scoring methods, or different outcomes.

ODARA risk categories approximated a flattened normal distribution: $M = 3.00$ ($SD = 1.69$), median = 3.00, skewness = 0.75, kurtosis = $-.916$. We created the following three treatment intensity categories, which allowed for the closest approximation of terciles without subdividing existing interpretation categories. ODARA scores from 0 to 2 were designated “Low” and comprised the lowest scoring 42% of the sample. “Medium” comprised ODARA scores 3 and 4, and 35% of the sample. “High” comprised ODARA scores 5–13 and the highest scoring 24% of the sample.

DISCUSSION

As noted previously, the intent of this study was to demonstrate how BIP service providers can use the ODARA and its scoring framework to establish low, medium, and high categories for their IPV offenders, and consequently use them to inform treatment intensity decisions (as outlined by the risk principle of the PEI). Results from this reanalysis of data from a subsample of the original ODARA construction and validation research found that the ODARA significantly predicted domestic violence recidivism among the 970 men who had been charged for the index assault. ODARA scores and corresponding risk categories were approximately normally distributed.

We set a rule of identifying categories for treatment intensity based on existing ODARA risk categories, in order to simplify the implementation of treatment intensity guidelines based on the present data, and to avoid discordance with the categories used for risk communication. We set a second rule based on percentile distributions in order to illustrate how BIPs can identify low, medium, and high treatment need groups based on ODARA scores alone. This rule means that BIPs can replicate our illustration in their own agency or state based on data easily obtainable from intake assessments with their own population, without the need to conduct their own follow-up research on domestic violence recidivism. However, the preponderance of cases in nonextreme risk categories, as indicated by percentiles, meant that the three treatment intensity categories could not be assigned equal proportions of the sample. We proceeded with the designation of the high, medium, and low categories based on a reasonable application of our rules, and hope that this straightforward procedure will assist BIP agencies in creating similar categories to guide the provision of treatment services based on the treatment intensity demands among their own clientele.

Clinical Implications

We have illustrated a procedure for deriving treatment intensity categories from a domestic violence risk assessment tool that has been widely adopted in criminal justice systems across the United States and Canada. Locally developed treatment need categories would facilitate BIPs making the most effective use of their actual treatment resources. By assessing their own clients and examining the resulting distribution of ODARA scores, BIPs could create their own treatment intensity categories using the procedure we have illustrated without having to do follow-up research to gather recidivism data in their service population. They could predetermine what portion of their clientele they can offer high-intensity treatment to, based on the resources available, and select treatment intensity categories that correspond to that proportion of the offenders referred to their program. In our illustration, we aimed for equal thirds and opted not to break up categories used to interpret recidivism risk, for reasons of convenience and ease of presentation in the present study, but BIPs need not adhere to these criteria. For example, if BIPs find that their clientele's average ODARA scores are higher than the sample we used, or can only provide high-intensity treatment to a small percentage of their referrals, they may choose to set their high-intensity category cutoff at an ODARA score of 8 or even higher. Thus, BIPs could offer their most intensive treatment services to the offenders who are highest need while managing their actual resources.

Treatment intensity categories are distinct from risk categories, which are used for the purpose of simplifying risk communication and are based on rates of recidivism. Recidivism estimates are best communicated numerically, along with percentiles and other information that aid interpretation of each individual offender's risk in comparison with normative data for the population (e.g., Hanson, Lloyd, Helmus, & Thornton, 2012; Scurich, 2018). Treatment intensity categories, on the other hand, are intended to facilitate the apportioning of treatment services, where different levels of treatment

duration, dosage, and content are available. Using the present illustration, risk and treatment intensity for an offender might be communicated as follows: “Mr. X scores 3 on the ODARA. As indicated by the normative data, 47% of men obtained higher scores than Mr. X. When followed up for approximately 5 years, 34% of men in Mr. X’s risk category committed a new domestic assault. This qualifies Mr. X for the Medium treatment intensity category for our service.” BIPs who examine the distribution of ODARA scores in their own referral population would refer to their own population data instead of the published normative data (Hilton et al., 2010). In this way, treatment intensity categories can guide implementation of the PEI, to the extent that the most intensive, evidence-based treatments can be given to domestic offenders with the highest priority for treatment, and given to them sooner where demand exceeds resources.

Assigning offenders to different treatment intensities is only one of several principles involved in providing effective correctional service (Bonta & Andres, 2016). The next step is to assess offenders’ criminogenic needs, such as antisocial personality traits, procriminal attitudes, substance use, and employment problems, that require attention during treatment (Hilton & Radatz, 2018). Offenders in the high-intensity category for treatment will also be a priority for further assessment to identify their criminogenic needs, as well as other characteristics that may impact their responsiveness to treatment, such as mental disorders.

Limitations and Research Implications

We reanalyzed a dataset in which ODARA scores were normally distributed and relatively few cases scored in the highest ODARA risk category. However, when the ODARA is scored prospectively rather than derived from research variables, higher mean scores have been reported compared to the original published norms (e.g., Jung & Buro, 2017). This finding could be attributable to the present sample dating from the 1990s, when proarrest policies in policing were relatively new (e.g., Hilton & Harris, 2009b) and IPV risk assessment was not a priority, whereas more recently police services have adopted a risk assessment approach (e.g., Hilton & Eke, 2017) such that risk-relevant information is routinely collected and may be more readily available in official reports, resulting in higher scores. Furthermore, we scored missing items as 0, whereas the published scoring instructions suggest prorating the ODARA score for items that appear to be present but for which documentation is unclear or incomplete (Hilton et al., 2010). Consequently, the proportion of men falling into the high treatment intensity category may be considered conservative. Further research that applies the ODARA prorating procedure where appropriate could be used to derive treatment intensity categories based on prorated scores where needed.

The present dataset included only men who had assaulted a woman with whom they were married or cohabiting at the time of the index assault or previous to it. Dating violence was not included, although the ODARA has since been validated in samples that included dating violence offenders (e.g., Hilton & Eke, 2016; Hilton, Popham, Lang, & Harris, 2014). BIP agencies that create treatment intensity categories using large samples of IPV offenders regardless of marital status could be encouraged to publish their research.

CONCLUSION

We explored how the ODARA, a tool developed to identify risk of IPV recidivism, can be used by BIPs to guide treatment intensity decisions for IPV offenders. We illustrated how the distribution of scores on the ODARA can be used to inform a three-tiered categorical system to match offenders to low, medium, and high levels of treatment, consistent with the risk principle of the PEI. We encourage BIPs to examine how treatment intensity categories could help them implement the risk principle with their own clientele. By using a tool such as the ODARA, BIPs could assess their clients' risk at intake and test category cutoffs that optimize their treatment resources by allocating their highest risk offenders to their most intensive treatment stream.

REFERENCES

- Andrews, D. A., & Bonta, J. (1995). *The level of service inventory—Revised (LSI-R) manual*. North Tonawanda, NY: Multi-Health Systems.
- Andrews, D. A., & Bonta, J. (2001). *LSI-R user's manual*. Toronto, ON: Multi-Health Systems.
- Andrews, D. A., Bonta, J., & Hoge, R. D. (1990). Classification for effective rehabilitation: Rediscovering psychology. *Criminal Justice and Behavior, 17*, 19–52. doi:10.1177/0093854890017001004
- Andrews, D. A., Zinger, I., Hoge, R. D., Bonta, J., Gendreau, P., & Cullen, F. T. (1990). Does correctional treatment work? A clinically relevant and psychologically informed meta-analysis. *Criminology, 28*(3), 369–404. doi:10.1111/j.1745-9125.1990.tb01330.x
- Babcock, J. C., Armenti, N., Cannon, C., Lauve-Moon, K., Buttell, F. P., Ferreira, R., . . . Solano, I. (2016). Domestic violence perpetrator programs: A proposal for evidence-based standards in the United States. *Partner Abuse, 7*(4), 355–460. doi:10.1891/1946-6560.7.4.355
- Belfrage, H., & Strand, S. (2012). Measuring the outcome of structured spousal violence risk assessments using the B-SAFER: Risk in relation to recidivism and intervention. *Behavioral Sciences & the Law, 30*, 420–430. doi:10.1002/bsl.2019
- Bonta, J., & Andrews, D. A. (2016). *The psychology of criminal conduct* (6th ed.). New York, NY: Routledge.
- Campbell, J. C., Webster, D. W., & Glass, N. (2009). The danger assessment: Validation of a lethality risk assessment instrument for intimate partner femicide. *Journal of Interpersonal Violence, 24*(4), 653–674. doi:10.1177/0886260508317180
- Cannon, C., Hamel, J., Buttell, F. P., & Ferreira, R. (2016). A survey of domestic violence perpetrator programs in the United States and Canada: Findings and implications for policy and intervention. *Partner Abuse, 7*(3), 226–276. doi:10.1891/1946-6560.7.3.226
- Connors, A. D., Mills, J. F., & Gray, A. L. (2012). An evaluation of intimate partner violence intervention with incarcerated offenders. *Journal of Interpersonal Violence, 27*(6), 1176–1196. doi:10.1177/0886260511424499
- Connors, A. D., Mills, J. F., & Gray, A. L. (2013). Intimate partner violence intervention for high-risk offenders. *Psychological Services, 10*(1), 12–23. doi:10.1037/a0028979
- Gerbrandij, J., Rosenfeld, B., Nijdam-Jones, A., & Galietta, M. (2018). Evaluating risk assessment instruments for intimate partner stalking and intimate partner violence. *Journal of Threat Assessment and Management, 5*, 103–118. doi:10.1037/tam0000101

- Gerth, J., Rossegger, A., Singh, J. P., & Endrass, J. (2015). Assessing the risk of severe intimate partner violence: Validating the DyRiAS in Switzerland. *Archives of Forensic Psychology, 1*(2), 1–15. doi:10.16927/afp.2015.1.1
- Gover, A. R. (2011). New directions for domestic violence offender treatment standards: Colorado's innovative approach to differentiated treatment. *Partner Abuse, 2*(1), 95–120. doi:10.1891/1946-6560.2.1.95
- Gover, A. R., Richards, T. N., & Tomsich, E. A. (2015). *Colorado's innovative response to domestic violence offender treatment: Current achievements and recommendations for the future*. Denver, CO: Buechner Institute for Governance. Retrieved from <https://cdpsdocs.state.co.us/dvomb/Research/UCDDV.pdf>
- Hansen, J., Davis, C., Smith, J. M., & Hilkey, S. (2016). *Standards for treatment with court ordered domestic violence offenders: A process evaluation*. Denver, CO: Colorado Division of Criminal Justice. Retrieved from <https://cdpsdocs.state.co.us/dvomb/Research/Evaluation.pdf>
- Hanson, R. K., Lloyd, C. D., Helmus, L., & Thornton, D. (2012). Developing non-arbitrary metrics for risk communication: Percentile ranks for the Static-99/R and Static-2002/R sexual offender risk tools. *International Journal of Forensic Mental Health, 11*, 9–23. doi:10.1080/14999013.2012.667511
- Hilton, N. Z., & Eke, A. W. (2016). Non-specialization of criminal careers among intimate partner violence offenders. *Criminal Justice and Behavior, 43*, 1347–1363. doi:10.1177/0093854816637886
- Hilton, N. Z., & Eke, A. W. (2017). Assessing risk of intimate partner violence. In J. C. Campbell & J. T. Messing (Eds.), *Assessing dangerousness: Domestic violence offenders and child abusers* (3rd ed., pp. 139–178). New York, NY: Springer Publishing.
- Hilton, N. Z., & Ham, E. (2015). Cost-effectiveness of electronic training in domestic violence risk assessment: ODARA 101. *Journal of Interpersonal Violence, 30*, 1065–1073. doi:10.1177/0886260514539762
- Hilton, N. Z., & Harris, G. T. (2009a). How nonrecidivism affects predictive accuracy: Evidence from a cross-validation of the Ontario Domestic Assault Risk Assessment (ODARA). *Journal of Interpersonal Violence, 24*, 326–337. doi:10.1177/0886260508316478
- Hilton, N. Z., & Harris, G. T. (2009b). Criminal justice responses to partner violence: History, evaluation, and lessons learned. In J. Lutzker & D. Whitaker (Eds.), *Preventing partner violence: Foundations, intervention, issues* (pp. 219–243). Washington, DC: American Psychological Association.
- Hilton, N. Z., Harris, G. T., & Rice, M. E. (2001). Predicting violence by serious wife assaulters. *Journal of Interpersonal Violence, 16*(5), 408–423. doi:10.1177/088626001016005002
- Hilton, N. Z., Harris, G. T., & Rice, M. E. (2010). *Risk assessment for domestically violent men: Tools for criminal justice, offender intervention, and victim services*. Washington, DC: American Psychological Association.
- Hilton, N. Z., Harris, G. T., Rice, M. E., Houghton, R., & Eke, A. W. (2008). An in-depth actuarial risk assessment for wife assault recidivism: The Domestic Violence Risk Appraisal Guide. *Law and Human Behavior, 32*, 150–163. doi:10.1007/s10979-007-9088-6
- Hilton, N. Z., Harris, G. T., Rice, M. E., Lang, C., Cormier, C.A., & Lines, K. J. (2004). A brief actuarial assessment for the prediction of wife assault recidivism: The Ontario domestic assault risk assessment. *Psychological Assessment, 16*, 267–275. doi:10.1037/1040-3590.16.3.267

- Hilton, N. Z., Popham, S., Lang, C., & Harris, G. T. (2014). Preliminary validation of the ODARA for female intimate partner violence offenders. *Partner Abuse, 5*, 189–203. doi:10.1891/1946-6560.5.2.189
- Hilton, N. Z., & Radatz, D. (2018). The criminogenic and noncriminogenic treatment needs of intimate partner violence offenders. *International Journal of Offender Therapy and Comparative Criminology, 62*(11), 3247–3259. doi:10.1177/0306624X17740015
- Jung, S., & Buro, K. (2017). Appraising risk for intimate partner violence in a police context. *Criminal Justice and Behavior, 44*, 240–260. doi:10.1177/0093854816667974
- Kropp, P. R. (2008). Intimate partner violence risk assessment and management. *Violence and Victims, 23*, 202–222. doi:10.1891/0886-6708.23.2.202
- Kropp, P. R., & Hart, S. D. (1997). *Assessing risk of violence in wife assaulters: The spousal assault risk assessment guide Impulsivity: Theory, assessment, and treatment* (pp. 302–325). New York, NY: Guilford Press.
- Kropp, P. R., Hart, S. D., Webster, C. D., & Eaves, D. (1998). *Spousal assault risk assessment: User's guide*. Toronto, ON: Multi-Health Systems.
- Kropp, P. R., Hart, S. D., Webster, C. D., & Eaves, D. (1999). *Spousal assault risk assessment guide*. Toronto, ON: Multi-Health Systems.
- Loinaz, I. (2014). Typologies, risk and recidivism in partner-violent men with the B-SAFER: A pilot study. *Psychology, Crime and Law, 20*, 183–198. doi:10.1080/1068316X.2013.770854
- Messing, J. T., & Thaller, J. (2013). The average predictive validity of intimate partner violence risk assessment instruments. *Journal of Interpersonal Violence, 28*, 1537–1558. doi:10.1177/0886260512468250
- Olver, M. E., & Jung, S. (2017). Incremental prediction of intimate partner violence: An examination of three risk measures. *Law and Human Behavior, 41*, 440–453. doi:10.1037/lhb0000251
- Radatz, D. L., & Wright, E. M. (2016). Integrating the principles of effective intervention into batterer intervention programming: The case for moving toward more evidence-based programming. *Trauma, Violence, & Abuse, 17*(1), 72–87. doi:10.1177/1524838014566695
- Rettenberger, M., & Eher, R. (2013). Actuarial risk assessment in sexually motivated intimate-partner violence. *Law and Human Behavior, 37*, 75–86. doi:10.1037/b0000001
- Rice, M. E., & Harris, G. T. (2005). Comparing effect sizes in follow-up studies: ROC, Cohen's d and r. *Law and Human Behavior, 29*, 615–620. doi:10.1007/s10979-005-6832-7
- Scurich, N. (2018). The case against categorical risk estimates. *Behavioral Sciences and the Law, 36*, 554–564.
- Stewart, L. A., Flight, J., & Slavin-Stewart, C. (2013). Applying effective corrections principles (RNR) to partner abuse interventions. *Partner Abuse, 4*(4), 494–534. doi:10.1891/1946-6560.4.4.494
- Stewart, L. A., Gabora, N., Kropp, P. R., & Lee, Z. (2014). Effectiveness of risk-needs-responsivity-based family violence programs with male offenders. *Journal of Family Violence, 29*(2), 151–164. doi:10.1007/s10896-013-9575-0
- Stewart, L. A., & Power, J. (2014). Profile and programming needs of federal offenders with histories of intimate partner violence. *Journal of Interpersonal Violence, 29*(15), 2723–2747. doi:10.1177/0886260514526059

- Storey, J. E., Kropp, P. R., Hart, S. D., Belfrage, H., & Strand, S. (2014). Assessment and management of risk for intimate partner violence by police officers using the Brief Spousal Assault Form for the Evaluation of Risk. *Criminal Justice & Behavior, 41*, 256–271. doi:10.1177/0093854813503960
- Williams, K. R., & Houghton, A. B. (2004). Assessing the risk of domestic violence reoffending: A validation study. *Law and Human Behavior, 28*(4), 437–455. doi:10.1023/B:LAHU.0000039334.59297.f0

Disclosure. N. Zoe Hilton is an author of the ODARA and declares a financial interest in a publication cited in this manuscript. Dana L. Radatz has no relevant financial interest or affiliations with any commercial interests related to the subjects discussed within this article.

Acknowledgments. Original data collection was supported by a grant from the Social Sciences and Humanities Research Council of Canada (SSHRC). We thank the police, probation, and archive services that contributed to the original data; members of the original research team Catherine Cormier, Carol Lang, Grant Harris, Kate Lines, Marnie Rice, and Ruth Houghton; and those who provided research and administrative assistance: Franco Bassiarello, Leslie Belchamber, Joseph Camilleri, Sonja Dey, Angela Eke, Marnie Foster, Julie McKay, Kelly Rawson, and Amilynn Sharpe. We also thank Angela Eke and Elke Ham for helpful comments on an earlier draft of this article.

Correspondence regarding this article should be directed to Dana L. Radatz, PhD, Niagara University, P.O. Box 1941, New York, NY 14109-2203. E-mail: dradatz@niagara.edu