An Indepth Actuarial Assessment for Wife Assault Recidivism:
The *Domestic Violence Risk Appraisal Guide*

N. Zoe Hilton, Grant T. Harris, Marnie E. Rice, Ruth E. Houghton and Angela W. Eke

How to cite TSpace items

UNIVERSITY OF TORONTO
LIBRARIES

Running Head: PREDICTING WIFE ASSAULT

An Indepth Actuarial Assessment for Wife Assault Recidivism:

The *Domestic Violence Risk Appraisal Guide*

Abstract

An actuarial tool, the *Ontario Domestic Assault Risk Assessment* (ODARA), predicts recidivism using only variables readily obtained by frontline police officers. Correctional settings permit more comprehensive assessments. In a subset of ODARA construction and cross-validation cases, 303 men with a police record for wife assault and a correctional system file, the VRAG, SARA, Danger Assessment, and DVSI also predicted recidivism, but the Hare Psychopathy Checklist (PCL-R) best improved prediction of recidivism, occurrence, frequency, severity, injury, and charges. In 346 new cases, ODARA and PCL-R independently predicted recidivism. An algorithm was derived for a combined instrument, the *Domestic Violence Risk Appraisal Guide* (DVRAG), and an experience table is presented (N = 649). Results indicated the importance of antisociality in wife assault.

An Indepth Actuarial Assessment for Wife Assault Recidivism:

The *Domestic Violence Risk Appraisal Guide*

As the 21st century began, the weight of empirical evidence indicated that assessment tasks are performed more accurately with the use of statistical rather than unstructured or informal methods. A meta-analysis by Grove and colleagues (Grove, Zald, Lebow, Snitz, & Nelson, 2000) found that formal methods (including statistical equations, actuarial tables, and algorithmic programs) were more accurate than clinical judgment (subjective, intuitive, and variable or unspecified processes for combining information) to render predictions. The advantage of statistical methods was consistent across time, whether assessors were psychologists, whether clinical information was available, and regardless of assessors' experience and access to information. Grove et al. (2000) reported that the size of this advantage varied with the outcome of interest, and was highest for forensic outcomes (e.g., inpatient assault, criminal recidivism). Ægisdóttir et al. (2006) updated the meta-analysis focusing on assessments pertinent to counseling psychology. They obtained similar results across a variety of statistical formulae, outcome criteria, and levels of clinician familiarity with the data. Violence risk assessments yielded the greatest superiority for statistical over informal methods.

Empirical support for the statistical prediction of violence raises optimism for similar performance in predicting wife assault specifically. Huss and Langhinrichsen-Rohling (2000) suggested that psychopaths represent a serious and persistent subgroup of wife assaulters (see also Johnson et al., 2006; Spidel et al., 2007). The Hare Psychopathy Checklist (PCL-R; Hare, 2003) is the accepted, standard forensic tool for measuring psychopathy and, although not designed as a risk assessment, its score has proven to be one of the most robust predictors of violent and criminal recidivism (Harris, Skilling, & Rice, 2001). It has the largest weight in an

actuarial tool for the assessment of the risk of violent recidivism, the *Violence Risk Appraisal*

Guide (VRAG, Harris, Rice, & Quinsey, 1993), which has over 30 published replications

(reviewed in Quinsey, Harris, Rice, & Cormier, 2006). The VRAG also predicted violent

recidivism by men with a history of severe wife assault or murder, better than the PCL-R alone

(Hilton, Harris, & Rice, 2001), and better than nonactuarial tools for the prediction of spousal

violence in a follow-up of wife assaulters (Grann & Wedin, 2002).

Domain-Specific Prediction

The *Ontario Domestic Assault Risk Assessment* (ODARA; Hilton et al., 2004) is an

actuarial risk assessment for wife assault recidivism that can be scored by frontline police

officers on the basis of readily available information. It contains thirteen empirically selected

items, some apparently specific to domestic relationships (prior domestic violence, confinement

of the victim, number of children, perpetrator assaulted victim when she was pregnant, victim's

children from prior relationships, victim's concern about future assaults, and barriers to victim

support), and several common to the literature on the risk of antisocial behavior in general, (prior

correctional sentence, failure on conditional release, substance abuse, threats of violence, and

two items pertaining to prior nondomestic violence). ODARA score predicted wife assault

recidivism in police records, with an ROC area of .77 in construction, and .72 in 100 cross-

validation cases. Because the ODARA was originally intended for use by police officers, only

information "routinely available in the field" was considered for inclusion (Hilton et al., 2004,

p.269), and information more difficult to obtain, such as PCL-R score, was not.

Hilton et al. (2004) reported that total score on the *Spousal Assault Risk Assessment*

(SARA), a nonactuarial assessment that can be scored from more indepth correctional and

clinical records (Kropp & Hart, 2000), predicted wife assault recidivism. The SARA can be

considered a candidate in the search for assessments to improve upon the information routinely available to frontline police officers. Similar methods were used to develop the *Domestic Violence Screening Instrument* (DVSI, Williams & Houghton, 2004) which also contains items requiring some clinical information (e.g., treatment participation) and the *Danger Assessment* (DA; Campbell, 1986; 1995) was associated with any new wife assault (Goodman, Dutton, & Bennet, 2000; Weisz, Tolman, & Saunders, 2000), although it is designed as a victim interview for predicting lethal assault in particular.

The present research addressed how the risk of wife assault recidivism is most accurately assessed when the more complete information typically available to forensic clinicians and criminal justice officials is available. Such assessors typically have access to extensive information about the perpetrator's history of criminal and antisocial behavior as well as victim reports and domain-specific information (e.g., relationship history) not typically available in front line police work, and should use the most accurate information available to assess risk. In this circumstance, use of a frontline risk assessment that does not take account of indepth clinical information about antisociality could lead to suboptimal prediction.

In correctional contexts, including those involving sentencing, parole, and community supervision, the concern is often not only the likelihood of violence, but also its severity. Detection of valid predictors might also be more powerful using continuous measures of recidivism (Heckert & Gondolf, 2004). Instruments designed to predict dichotomous outcomes (i.e., violent recidivism vs. not) might not predict violence severity or such other continuous outcomes as the number of recidivistic incidents, injury caused, severity of violent behavior, or the seriousness of charges incurred (though see Koziol-McLain et al., 2006). The present study, therefore, examined whether the prediction of dichotomous recidivism plus several indices of the

severity of recidivism could be enhanced by adding indepth material to the easy-to-gather ODARA items.

In addition, correctional populations might be higher risk than the entire population of offenders, and have less opportunity to recidivate than offenders who receive only community dispositions. We anticipated that incarceration would result in attenuated predictive accuracy of the instruments previously tested by Hilton et al. (2004). In recent research (Popham & Hilton, 2006), convicted domestic violence perpetrators were more likely to fall in the highest ODARA risk category than the population of perpetrators with a police record, of which they are a subgroup, raising the need for greater discrimination in this subpopulation.

The Present Study

We examined whether the prediction of wife assault recidivism and its severity could be improved by adding more detailed clinical information to the ODARA. Seto (2005) reported that predicting sex offenders' recidivism could not be improved by combining the most accurate actuarial risk assessment with scores from other actuarial tools; however, only tools specifically designed to predict recidivism among sex offenders were considered. We tested tools designed specifically for domestic violence (the SARA, DA, and DVSI), and indepth assessments known to predict violence in general (PCL-R and VRAG). We began with a sample of men with a police record of assault against a female cohabiting partner or ex-partner, from the initial ODARA construction and cross-validation, but used only those cases that had a more detailed correctional system case file (usually a probation or pre-sentence report). Results from this sample were tested in a new sample of cases (not used in any previous development or validation) drawn from three police sources, with additional clinical material from correctional files.

Method

Inclusion Criteria, Data Bases, and Selection

From the three police records management systems described below, we identified each male perpetrator in incidents classified as domestic. For each perpetrator, we then isolated the incident closest to, but no later than, December 31, 1996, (Sample 1) or 1997 (Sample 2) in which, according to a victim/witness report and other police evidence, he committed an act of physical assault or credible threat of death with weapon in hand in the presence of a victim who was a current or former wife or common-law wife (hereafter called the index assault). Eligible cases were all those with evidence of both an intimate relationship and an existing or prior marital or cohabiting relationship. Incidents involving only nonspousal victims or nonviolent acts were not eligible as the index incident, and cases in which offenders and victims had not lived together were also excluded, mainly because it was not always discernible from police reports of assaults involving noncohabiting relationships whether the relationships were intimate. Eligibility did not require that offenders be arrested or criminally charged for the index assault, but research assistants only coded cases where they could agree, based on the police reports, that an assault had clearly occurred. In the case of multiple eligible incidents, the one closest to the December 31 cut-off was coded as the index assault.

Cases for Sample 1 were drawn from an electronic records management system used by the Ontario Provincial Police as well as by approximately 50 municipal police services. Most rural areas, Aboriginal communities, and many municipalities in Ontario, Canada's most populous province, were represented in this archive. The records included dispatch information and verbatim reports by frontline officers; names of the perpetrator (suspect), victims, and witnesses; charges laid; and all details of the investigation including statements by all parties. We were provided a list of all entries classified as domestic, and research assistants retrieved the

records and determined eligibility. We began with 689 cases (used for the construction and cross-validation of the ODARA; Hilton et al., 2004) from which we selected all those with files compiled by the Ministry of Corrections (e.g., pre-sentence reports, psychological evaluations, psychosocial assessments, probation officer notes, etc.) as a result of charges pertaining to the index or any other offense. The existence and location of these files was ascertained from a Ministry offender tracking data base. This selection process yielded a sample of 303 men with a police report of wife assault all of whom had a corrections file containing indepth clinical and psychosocial information. The other cases from the 689 did not have a correctional file and are not used in the present study. Most index assaults occurred in 1996 (74%) or 1995 (19%); 75% were recorded as having resulted in criminal charges.

For Sample 2 we obtained new cases, not used in Sample 1 or any previous study, using the same identification and selection process as in Sample 1. Approximately half (n = 168, 49%) were retrieved from the same source as used to compile Sample 1. The rest (n = 178, 51%) were from similar records management systems maintained by two urban police services in the Greater Toronto Area. All cases were selected first from the police archives for eligibility and then from the Ministry offender tracking database for having a corrections file. In this sample of 346 cases, most index assaults occurred in 1996 (71%) or 1997 (17%); 90% resulted in criminal charges (possibly a result of greater emphasis on charges in the metropolitan jurisdictions).

For each perpetrator in Samples 1 and 2, we supplemented information on pre-index criminal history, index disposition, and recidivism with criminal record data obtained from a federal database maintained the Royal Canadian Mounted Police, which records all criminal charges, convictions, and criminal dispositions across Canada.

Procedure and Variables

All file information was retrieved within, or transported securely to, a secure location in a police research unit. Sociodemographic variables, criminal history, items from formal assessments, and outcomes were coded at the same time using all documentation available regarding the index and prior history. Coding was done by researchers with many years' experience coding such information, who also trained and closely supervised two graduate assistants with one to three years' coding experience. We created a manual for quantifying and coding all data available in the police reports plus information in corrections files about childhood history, adult functioning, relationships, and assessment. Variables describing the two samples are shown in Table 1; those requiring fuller explanations are described next.

Substance abuse score. As per scoring instructions for this ODARA item, 1 point is given for each of eight variables that predicted violent recidivism in previous research (Harris et al., 1993): perpetrator used alcohol immediately before or during the index assault, perpetrator used drugs immediately before or during the index assault, perpetrator abused alcohol or drugs in the days or weeks (up to a month) leading up to the index date, perpetrator used alcohol or drugs more than usual in the days or weeks (up to a month) leading up to the index date, perpetrator is more angry or violent when using alcohol or drugs, perpetrator was previously charged for an offense while under the effects of alcohol, perpetrator had an alcohol problem since age 18, and perpetrator had a drug problem since age 18.

Prior domestic incidents. We recorded all police reports pertaining to incidents, separate from and occurring before the index assault, involving a forceful physical contact or threat of physical harm by the perpetrator against the victim of the index assault or a previous female partner with whom he had lived or was living, or the partner's child(ren). Similar information was recorded for prior nondomestic incidents (i.e., violence against any other persons).

Prior criminal history score. We used the Cormier–Lang Criminal History score (Quinsey et al., 2006), which captures the frequency and severity of criminal conduct by totaling all charges ranging from 1 (minor property offense) to 28 (homicide).

Violation of prior conditional release order. We considered any occasion on which the perpetrator, while living in the community, disobeyed any order of a criminal court, or civil court in the case of no-contact orders. Examples of violation include: committing a new criminal offense while on bail, probation, parole, or suspended sentence; failure to appear for court or appointment with a parole or probation officer; drinking or having firearms when prohibited; contacting a person when prohibited. We did not require criminal charges for the violation.

Injury. For all injury measures, we used a 5-point ordinal scale derived from the Danger Assessment (Campbell, 1985, p.105): 1 (no injuries or lasting pain), 2 (bruises, cuts, or continuing pain), 3 (severe contusions, burns, broken bones), 4 (head injury, internal injury, or permanent injury), 5 (wounds from a weapon; e.g., stabbed, shot).

Severe violence. We recorded whether the perpetrator committed acts defined as severe on the *Revised Conflict Tactics Scales* (CTS-2; Straus, Hamby, Boney-McCoy, & Sugarman, 1996): punched, kicked, bit, hit with something that could hurt, beat up, choked, slammed against wall, burned or scalded, used a knife or gun; plus similarly severe acts.

Violence Risk Appraisal Guide (VRAG). The VRAG (Quinsey et al., 2006) was developed for male offenders (Harris et al., 1993) and predicts violent recidivism among forensic patients (Harris, Rice, & Cormier, 2002), nonforensic psychiatric patients (Harris, Rice, & Camilleri, 2004), sex offenders (Harris et al., 2003), and released prisoners (Glover, Nicholson, Hemmati, Bernfeld, & Quinsey, 2002). In construction, its 12 items, including PCL-R score, all significantly and uniquely predicted violent recidivism (Harris et al. 1993). The VRAG has

repeatedly achieved inter-rater reliability coefficients above .90 (Quinsey et al., 2006). In the

present study, some items had no variance (diagnosis of schizophrenia, female victim, and ever

married or equivalent), reducing the VRAG to a 9-item modification. Previous modifications

replicated the VRAG's accuracy, albeit with attenuated effect sizes (see Harris & Rice, 2003).

Hare Psychopathy Checklist – Revised (PCL-R) score. The 20-item PCL-R (Hare, 2003)

is the standard tool for forensic assessment of psychopathy. Although scoring is often based on

interview plus file information (Hare, 2003), we used file information alone, which yields

reliable and valid ratings, especially for violence prediction (Hare, 2003; Harris et al., 2001).

The Spousal Assault Risk Assessment (SARA). The SARA (Kropp, Hart, Webster, &

Eaves, 1999) is a structured professional judgment scheme for domestic violence. Its 20 items,

gleaned from empirical and clinical literatures, are scored 0, 1, or 2. All 20 items were scored

and summed for an item total score. Although the SARA manual advises that interviews with the

accused perpetrator and victim(s) be conducted, the total score coded from file information has

achieved inter-rater reliability over .80 (Kropp & Hart, 2000) and has predicted wife assault

recidivism (Grann & Wedin, 2002; Heckert & Gondolf, 2004; Hilton et al., 2004).

Danger Assessment (DA). The DA (Campbell, 1986, 1995) is an abuse history interview

plus a structured scale designed to assess the risk of lethal wife assault. The items cover the

offender's domestic and nondomestic violence history, access to weapons, substance abuse,

jealousy, sexual assault, threats, and the victim's fear of being killed. Test-retest reliability has

been at or above .89 in interviews (Campbell, 1995). We coded all 15 items and summed them

for an item score total. This unweighted score has been reported to predict wife assault

recidivism (e.g., Goodman, Dutton, & Bennett, 2000; Heckert & Gondolf, 2004; Hilton et al.,

2004; Weisz, Tolman, & Saunders, 2000). In the present study, we deviated from the intended

implementation of the DA by scoring it from file material rather than a victim interview.

Domestic Violence Screening Instrument (DVSI). The DVSI (Williams & Houghton,

2004) contains 12 items derived from a 34-item clinical guide collected on over 9000 cases

analyzed for characteristics associated with men's history of repeated domestic violence,

supplemented by literature reviews and focus groups to identify items deemed to predict such

violence. Each item has two to four possible scores and the total score is the sum of the item

scores. In the development of the DVSI, this score had an internal consistency of .71 (inter-rater

reliability was not reported) and predicted severe wife assault recidivism, ROC = .68 (Williams

& Houghton, 2004). The DVSI-R, in which missing information was scored as 0, distinguished

between first-time and repeat offenders being assessed for violence towards partners or children,

ROC = .71 (Williams & Grant, 2006). The DVSI-R also added a structured clinical judgment of

low, moderate or high risk of violence, which was a significantly worse predictor than raw item

total score (Williams & Grant, 2006) and was not used in the present study. The DVSI was

designed to be scored from documentary material. Because it was published after our data were

collected, we scored it from other variables. Missing information was coded 0 as in the DVSI-R.

Two items (domestic violence treatment, history of domestic violence restraining orders) could

only be scored dichotomously rather than trichotomously. One item (drug or alcohol treatment)

was missing for all cases, making our implementation an 11-item modification.

Wife assault recidivism. Information about criminal and assaultive behavior occurring

after the index assault was obtained from police, corrections, and criminal record reports, up to

the end of 2001, a mean of 5.10 years (*SD* = 1.44) post-index. Subsequent assaults against a

current or former wife or common-law wife were deemed wife assault recidivism. We recorded

whether any incident occurred (dichotomous wife assault recidivism), the number of such

assaults in the follow-up period, the number of incidents with severe violence, and the Cormier-

Lang score for all criminal offenses associated with wife assault recidivism. Arrest, charges, or

convictions were not required to count as recidivism.

     Reliability. Inter-rater reliability was measured by two research assistants independently

scoring a random subsample of 30 cases (previously reported by Hilton et al., 2004). Individual

variables, including those comprising the ODARA, for which agreement met or exceeded a

Pearson correlation coefficient of .80 (continuous measures) or kappa coefficient of .70

(categorical variables) in that test were retained. Subsequently, ODARA score in 24 cases coded

by experienced coders, 10 of which were in addition scored by completely novice assessors, all

masked for outcome, achieved interrater reliabilities $\geq$ .90. Dichotomous recidivism coded

independently by two research assistants with all other information masked, also achieved an

inter-rater reliability > .90 (Hilton et al., 2004).

<div align="center">Results</div>

Sample Characteristics and Recidivism

     Table 1 shows the characteristics of Samples 1 and 2. The samples differed on some

variables (Table 1), and Sample 1 had significantly higher scores on three of the formal

assessments. Although a subset of the original ODARA construction and cross-validation cases,

Sample 1 used only cases with a corrections file and thus represents a higher risk group than

previously published. As a result, the predictive accuracy of the ODARA in this restricted-range

sample (recidivism base rate 49%) was marginally lower than in the unselected, previous cross-

validation sample, r (303) = .30, p < .001, ROC area = .67, 95% CI = .61-.73, suggesting there is

benefit in developing an improved prediction tool for a correctional sample.

Combining Assessments to Improve Predictive Accuracy

As shown in Table 2, all assessments considered as potential additions to the ODARA in Sample 1 were significantly and positively associated with dichotomous wife assault recidivism (i.e., at least one post-index assault against a current or former wife or common-law wife, documented in a police report). Each assessment was also significantly correlated with the ODARA in Sample 1: VRAG $r = .52$; PCL-R, $r = .57$; SARA, $r = .57$; DA, $r = .40$; DVSI, $r = .49$; all $p < .001$. Table 2 also shows the correlation between each candidate assessment and each of the continuous measures of recidivism. For each measure of recidivism, we conducted a binary logistic (for dichotomous recidivism) or linear (for continuous measures) regression[1] analysis for each formal assessment paired with the ODARA. No formal assessment made an incremental and independent improvement on the ODARA for dichotomous recidivism (all $p > .10$). For each of the continuous measures of recidivism, only one formal assessment significantly improved on the predictive accuracy of the ODARA. For the number of recidivistic incidents, adding PCL-R scores yielded a multiple $R = .28$, $\beta = .13$, $F(1, 300) = 4.05$, $p < .05$. For number of incidents with severe violence, adding PCL-R scores yielded multiple $R = .26$, $\beta = .13$, $F(1, 282) = 3.75$, $p < .06$; and for total recidivism injury, adding PCL-R scores yielded multiple $R = .31$, $\beta = .22$, $F(1, 300) = 9.36$, $p < .01$. Only adding DVSI scores to the ODARA improved the prediction of Cormier-Lang Criminal History score totaled across all recidivistic incidents, multiple $R = .26$, $\beta = .16$, $F(1, 300) = 6.88$, $p < .01$. Because the PCL-R was the most consistent contributor to the prediction of continuous outcomes, was the best bivariate predictor, and had yielded high levels of predictive accuracy in separate previous studies of serious domestic violence perpetrators (Hilton et al., 2001; Grann & Wedin, 2002), we selected it as

most likely to improve upon the ODARA in the prediction of wife assault recidivism with indepth clinical and psychosocial information.

The DVRAG, an Indepth Assessment for Wife Assault Recidivism

In order to optimize predictive accuracy and parsimonious combination of assessments given adequate time and information sources, the ODARA plus PCL-R score was identified as the basis of an indepth risk assessment. We used the Nuffield weighting system to derive individual weights for the 14 items (as used in the development of the VRAG; Harris et al., 1993; Quinsey et al., 2006), whereby weights are assigned to each value range of a variable according to its deviation from the base rate. For each 5% (rounded) deviation from the sample baserate in the recidivism rate among offenders with a given value range, a weight of plus or minus one is given to that value range. For six of the seven dichotomous variables in Sample 1 (Prior conditional release failure, Threat, Confinement, Victim concern, Violence against others, Assault when pregnant) a score of zero had a weight of 0. Prior correctional sentence was the exception: its absence yielded a weight of –2 because perpetrators scoring zero had a recidivism rate approximately 10% below the base rate. Offenders who made a Threat were 3% more likely than the base rate to recidivate, which was rounded to 5% for a weight of +1. All item weights were summed for each perpetrator, creating an assessment (see Appendix) whose scores reflect deviation from the base rate of wife assault recidivism (49%) in Sample 1.

The name *Domestic Violence Risk Appraisal Guide* (DVRAG) reflected the similar development of this new assessment to that of the VRAG. DVRAG scores were highly related to each outcome variable in Sample 1: Pearson *r* (n = 303) = .38, *p* < .001 for dichotomous wife assault recidivism; .40, *p* < .001 for the number of recidivistic offenses; .37, p < .001 for CTS severe domestic violence recidivism; .39, *p* < .001 for total victim injury in recidivistic offenses;

and .33, $p < .001$ for Cormier-Lang score. For each correlation, DVRAG score was a statistically

significant improvement over ODARA score (z score of difference > 1.65; Kanji, 1993), and

DVRAG score remained a significant predictor of each outcome measure when controlling for

ODARA score, rs .19 to .31, ps $< .01$. The ROC area for the predictive ability of DVRAG scores

in Sample 1 was .71, $SE = .03$, corresponding to a Cohen's $d = .80$ (Rice & Harris, 2005), and

represented a statistically significant improvement ($p < .05$, 1-tailed) over ODARA scores alone.

The 346 cases in Sample 2 were used to cross-validate both the ODARA and DVRAG,

because the cases were not previously used to construct or validate either assessment. The

predictive validity (recidivism base rate = 41%) of the ODARA was ROC area = .65 ($SE = .03$,

95% CI = .59 to .71), $d = .55$; and of the DVRAG, ROC area = .70 ($SE = .03$, 95% CI = .64 to

.75), $d = .75$; DVRAG score represented an improvement in predictive validity over ODARA

score, $p < .05$, 1-tailed (Sample 2).[2] Table 3 shows the correlation between each of the ODARA

and DVRAG and each of the continuous outcome variables in Sample 2. DVRAG correlations

remained significantly associated with each outcome when controlling for the ODARA, and

were significantly larger than those for the ODARA (z score of difference > 1.65) for

dichotomous wife assault recidivism, the number of recidivist incidents, and their severity in

terms of injury cases. The nonsignificant improvement in predicting Cormier-Lang score for

recidivism (which is a count of criminal charges weighted for severity), and the lower predictive

accuracy for this outcome compared to measures not requiring criminal charges, suggests that the

DVRAG predicts actual reported incidents rather than just police arrest decisions (which are

themselves influenced by many factors; Hilton, Harris, & Rice, in press). These results represent

successful replications for both the ODARA and DVRAG using cases not included in the

development or other testing of either instrument.

We combined the two samples to develop an actuarial experience table for the DVRAG. In the combined sample of 649 cases, mean DVRAG score was $M = 2.88$ ($SD = 7.76$), with a range of -10 to +37, and significant predictive accuracy for wife assault recidivism (base rate = 45%), ROC area = .70 (SE = .02, 95% CI $\pm$.04), $p < .001$, $d = .75$. The top scoring 3% of cases on the DVRAG were all recidivists. The distribution of scores also permitted identification of a category containing approximately 3% of the lowest scoring cases, and subdivision of remaining scores into seven equal sized categories at approximately the 20th, 40th, 60th, and 80th percentiles, which yielded a statistically reliable increase in recidivism rates at each score cut-off (Table 4). We evaluated the inter-rater reliability of DVRAG scores by having two independent raters score the DVRAG (blind to recidivism and each other's scores) using 10 cases randomly selected from the case files of a correctional treatment program for domestic violence perpetrators. The Pearson correlation coefficient was .92, and the intra-class correlation (absolute agreement, random effects, single measures) was .90, both $p < .001$. This level of reliability for the DVRAG and its overall standard deviation yielded a standard error of measurement of 2.2 (Nunnally, 1959). As the DVRAG categories are no smaller than four points each (Table 4), a perpetrator's obtained score is likely to be misclassified by no more than one category, 95% of the time. Finally, we evaluated the scoring reliability of the DVRAG by comparing the scoring by a research assistant to that of an experienced forensic clinician with no explicit DVRAG training, just the guidelines shown in the Appendix, for 16 cases. The resulting Pearson correlation was .83, indicating that the DVRAG is likely to be a reliable tool in regular forensic practices.

Direct comparisons of the predictive performance of the DVRAG and ODARA with the other formal assessments in the combined samples might be biased for several reasons: almost half of the 649 cases were used to construct each of the ODARA and DVRAG; some

compromises from the recommended scoring were required (e.g., interviews are usual for the

PCL-R, SARA, and DA); some DVSI items were unavailable; and some VRAG items were

invariant. Yet, all showed statistically significant predictive validity (Table 5). It is notable that

the VRAG and PCL-R performed at least as well as assessments designed specifically for

domestic violence risk assessment. This shared ability to predict recidivism to at least some

extent is not surprising given the inter-correlations among the assessments (Table 6), which in

turn is partly attributable to some overlapping items (e.g., nondomestic violence, conditional

release failure). On the other hand, different predictive accuracy may be an inevitable

consequence of some unique items (e.g., treatment history, confinement of the victim).[3]

Because the five outcome measures yielded similar levels of predictive accuracy, and

were predicted by similar variables, we computed the inter-correlations among all the measures

of wife assault recidivism. All were significantly associated with each other, with correlations

ranging from .36 (total injury and Cormier Lang score) to .75 (dichotomous wife assault and

number of severe incidents), all $p < .01$; however, they shared much less than half (on average,

30%) of their total common variance.

Discussion

Previous research had resulted in a simple frontline actuarial domestic violence risk

assessment for use by police officers, victim service workers, and bail courts, the Ontario

Domestic Assault Risk Assessment (ODARA; Hilton et al., 2004), whose 13 dichotomous items

rely on easily gathered information. The present study examined whether this simple tool could

be improved upon by adding tools that require indepth psychosocial and clinical information

usually available to clinicians and forensic professionals, including several formal assessments

for domestic violence (*Spousal Assault Risk Assessment*, SARA; Kropp & Hart, 2000; *Danger*

Assessment, DA; Campbell, 1995; *Domestic Violence Screening Instrument*, DVSI; Williams &

Houghton, 2004), general violence (*Violence Risk Appraisal Guide*, VRAG; Harris et al., 1993;

Quinsey et al., 2006), and lifecourse antisociality (*Hare Psychopathy Checklist*, PCL-R; Hare,

2003). In this sample of 649 cases with a correctional file, all assessments showed a significant

ability to predict wife assault recidivism as operationalized by at least some of several inter-

related outcome measures. Among these tools, however, the PCL-R showed the most promise in

improving prediction over the ODARA alone. Multivariate analyses in the development sample

(Sample 1) supported this finding.

The 14-item *Domestic Violence Risk Appraisal Guide* (DVRAG) comprises the original

ODARA items (but scored continuously rather than dichotomously) combined with the PCL-R.

Item weights were based on each constituent variable's empirical association with dichotomous

wife assault recidivism in a construction sample of perpetrators with a police report of a

domestic assault and a corrections file. Several tests on cases masked for outcome suggested that

the DVRAG can be expected to have good inter-rater reliability in routine forensic practice. Its

predictive accuracy corresponded to a "large" effect (where "large" is a Cohen's $d \geq .8$, ROC

area = .71, and "medium" is $d \geq .5$, ROC area = .64; Rice & Harris, 2005) in predicting wife

assault recidivism in the development sample of 303 cases, and its predictive accuracy was

maintained in a cross-validation sample of 346 cases and the combined sample of 649 cases. The

DVRAG performed better than the ODARA and the other formal assessments. No further

improvement was made by adding any other formal assessment or individual variable pertaining

to childhood abuse history, juvenile delinquency, and adult adjustment. The combined sample of

649 subjects was used to compile an actuarial experience table and a practical scoring scheme

(Appendix). Together, the ODARA and DVRAG represent not a proliferation of instruments but

a coherent system of risk assessment. For example, a police officer can score the ODARA in time for a bail decision and a forensic clinician or probation officer can subsequently score the DVRAG to provide an improved assessment to aid sentencing, supervision, and treatment decisions.

Assessing the Risk of Wife Assault Recidivism

The present results regarding optimal indepth assessment of wife assault risk closely parallel construction of the VRAG, which consists of several well established violence risk factors empirically combined with the PCL-R. The VRAG has yielded a large average effect in independent replications (Quinsey et al., 2006; www.mhcp-research.com/ragreps), especially when scored with high reliability and without modification (Harris & Rice, 2003). Together, this research confirms the value of actuarial methods for combining variables to construct decision support tools, especially risk assessments. Three published domestic violence risk assessments tested in the present study were not developed actuarially, but relied on existing domestic violence literature and clinical experience, neither of which has sufficiently taken account of criminogenic variables in previous violence prediction research (Hilton & Harris, 2005; Hilton & Simmons, 2001; Quinsey et al., 2006). In addition, items with the heaviest weights in the DVRAG tend to be those that reflect general antisociality. Thus, rational selection and combination of items appeared to have inadvertently concentrated too much on domain-specific characteristics (e.g., jealousy) when measures associated with and predictive of general criminality and violence (e.g., non-domestic criminal history) would have performed as well or better in predicting domestic violence recidivism.

This is not to say that the nonactuarial assessments did not predict wife assault recidivism. Indeed, we believe the present results constitute a successful replication of SARA

total scores, provide further evidence that DA scores predict wife assault recidivism, and are the first independent replication of the DVSI. On the other hand, we observed less variance in the prevalence of wife assault recidivism attributable to variables specific to domestic relationships than to enduring general antisociality. This observation might be qualified by the discovery of domain-specific variables that account for wife assault but were not adequately measured in the present study; however, the inability of three domestic violence risk assessments, developed using constructs identified in the 30-year-old domestic violence literature, to outperform either the VRAG or the PCL-R, leaves little optimism that such variables will be found.

Most of the tested assessments predicted wife assault recidivism across all outcome measures (dichotomous recidivism, number of incidents, total injury score, incidents with severe violence, and Cormier-Lang score), contrary to some expectations that different outcomes require different predictors. The outcomes were significantly intercorrelated but were not highly collinear. We infer that this pattern of results reflects considerable amount of measurement error in the evaluation of wife assault recidivism. Such error could increase the difficulty of predicting the specific behavior of wife assault, compared with predicting violence in general (a similar observation has been made with respect to sexually violent recidivism; Rice, Harris, Lang, & Cormier, 2006). Our use of several outcome measures allowed us to replicate the predictive ability of previously published assessments using new outcome measures. Also, we were able to select, for the DVRAG, the assessment (PCL-R) that most consistently and robustly added to the predictive ability of the existing actuarial ODARA across these outcomes, increasing confidence in the DVRAG's applicability to correctional and forensic clinical settings.

The superior prediction yielded by assessments not designed for domestic violence risk assessment confirms a previous finding that the VRAG and PCL-R outperformed the SARA in

predicting wife assault recidivism (Grann & Wedin, 2002). These results are consistent with the hypothesis that attitudes and actions specific to domestic relationships play a minor etiologic role in wife assault. In contrast, enduring antisociality – a fundamental, cross-situational, and qualitatively distinct characteristic – is responsible for a disproportionate amount of violent behavior, including wife assault. Huss and Langhinrichsen-Rohling (2006) have also reported that a measure of psychopathy was associated with subtypes of men, within a clinical treatment sample, responsible for the most extensive violence against partners and other victims, even though the psychopathy scores of their sample were limited in range. We speculate that some apparently domain-specific variables might actually be triggers for the occurrence of a violent act. For example, high levels of psychopathy best predict *who* is likely to be a repetitive wife assaulter, but a tendency towards sexual jealousy and proprietariness might act as a proximal trigger, predicting *when* such violence is most likely. Some apparently domain-specific variables on the ODARA might also be proxies for general antisociality, such as prior domestic assaults as well as nondomestic (indicating criminal versatility) and assault during pregnancy (indicating callousness). This theoretical perspective requires development and testing in further research.

Limitations

Although the DVRAG reliably rank ordered wife assaulters with respect to their risk of wife assault recidivism, the actuarial experience tables cannot automatically be applied to new populations without knowledge of the base rate of recidivism in those populations (Harris & Rice, in press; Mossman, 2006). The present offenders all had a correctional file and were more likely to have been charged than the unselected sample of men with a police record of wife assault used to construct the ODARA. The present probability estimates might over-estimate the rates of police detected recidivism in other populations, especially in the lower categories. For

example, the lowest DVRAG category has a probability of recidivism of 14% (and a wide

confidence interval) compared with only 5% in the lowest ODARA category. The experience

tables are less likely to over-estimate risk for wife assaulters in the correctional system, and we

believe that the most appropriate application of the DVRAG is to such a population – for whom

indepth clinical assessment information is available.

The present study was, nevertheless, limited by its information sources. Criminal justice

(police and correctional) archives do not permit optimal assessment of attitudes and emotional

traits and could not, in the present study, support the coding of psychiatric symptoms. Structured

interviews and psychological assessment of the perpetrator might improve the measurement of

these characteristics and, thereby, the predictive accuracy of all the formal assessments in this

study. This limitation would apply as much to the PCL-R as to any of the other existing

assessments, yet it was a strong predictor. There is no basis to expect additional clinical material

to alter the conclusion that the PCL-R was the best addition to the ODARA in predicting wife

assault recidivism. We did have several sources of information to measure recidivism (criminal

record, institutional and community corrections reports, and police reports that included victim

statements); however, only officially documented behavior was defined as recidivism. This

definition maximizes the utility of the DVRAG for criminal justice contexts. It might, though,

under-represent the total number of wife assault occurrences.

The predictive accuracy of all assessments might have been limited by some restriction of

the opportunity for recidivism relative to previous research that included some cases with no

corrections record or even charges for the index incident (Hilton et al., 2004). More cases were

incarcerated for wife assault or other violence than in the ODARA construction sample. Wife

assault recidivism was not eliminated by this intervention, but coincident with higher scores on

all the formal assessments, it likely caused some attenuation of predictive performance in the available follow-up period. Two assessments (the VRAG and DVSI) might also have been hampered by scoring modifications. The VRAG exhibited reduced variability compared with populations of generally violent offenders (Quinsey et al., 2006). The DVSI, though not scored exactly as intended, compared favorably with the other domestic violence tools, especially in predicting criminal charges for recidivism. For the PCL-R, SARA, and the DA, administration procedures were altered by coding from files without conducting interviews. The DA is intended to assess risk of lethal assaults which we did not test, although its best performance was for injury, a necessary condition for lethal assault.

In conclusion, we present the Domestic Violence Risk Appraisal Guide, DVRAG, a new actuarial indepth wife assault risk assessment for wife assault recidivism, including scoring details and experience table. DVRAG scores exhibited good inter-rater reliability, and large, cross-validated effects in the prediction of several related outcomes reflecting the occurrence, frequency, and severity of wife assault recidivism. Prospective replication of the DVRAG scored entirely masked for recidivism is desirable, although previous research has established that its components (ODARA and PCL-R) exhibit predictive accuracy when scored by assessors masked to outcome (Hare, 2003; Hilton et al., 2004). The success of the ODARA and PCL-R in predicting wife assault recidivism, and the predictive value of their statistical combination, illustrates the power of empirical methods in the construction of assessments for forensic professionals, and the robustness of measures of antisociality in predicting domain-specific violence.

References

Ægisdóttir, S., White, M. J., Spengler, P. M., Maugherman, A. S., Anderson, L. A., et al. (2006). The meta-analysis of clinical judgment project: Fifty-six years of accumulated research on clinical versus statistical prediction. *The Counseling Psychologist, 34*.

Campbell, J. C. (1986). Nursing assessment for risk of homicide with battered women. *Advances in Nursing Science, 8,* 36-51.

Campbell, J. C. (1995). Prediction of homicide of and by battered women. In J. C. Campbell (Ed.), *Assessing dangerousness* (pp. 96-113). Thousand Oaks, CA: Sage.

Glover, A. J. J., Nicholson, D. E., Hemmati, T., Bernfeld, G. A., & Quinsey, V. L. (2002). A comparison of predictors of general and violent recidivism among high risk federal offenders. *Criminal Justice and Behavior, 29*, 235-249.

Goodman, L. A., Dutton, M. A., & Bennett, L. (2000). Predicting repeat abuse among arrested batterers. *Journal of Interpersonal Violence, 15,* 63-74.

Grann, M., & Wedin, I. (2002). Risk factors for recidivism among spousal assault and spousal homicide offenders. *Psychology, Crime, and Law, 8,* 5-23.

Grove, W. M., Zald, D. H., Lebow, B. S., Snitz, B. E., & Nelson, C. (2000). Clinical versus mechanical prediction: A meta-analysis. *Psychological Assessment, 12,* 19-30.

Hare, R. D. (2003). *Hare Psychopathy Checklist-Revised (2nd ed.).* Toronto, ON: Multi-Health Systems Inc.

Harris, G. T., & Rice, M. E. (2003). Actuarial assessment of risk among sex offenders. *Annals of the New York Academy of Sciences, 989,* 198-210.

Harris, G. T. & Rice, M. E. (in press). Characterizing the value of actuarial violence risk assessment. *Criminal Justice and Behavior*.

Harris, G. T., Rice, M. E., & Camilleri, J. A. (2004). Applying a forensic actuarial assessment (the Violence Risk Appraisal Guide) to nonforensic patients. *Journal of Interpersonal Violence, 19,* 1063-1074.

Harris, G. T., Rice, M. E., & Cormier, C. A. (2002). Prospective replication of the Violence Risk Appraisal Guide in predicting violent recidivism among forensic patients. *Law and Human Behavior, 26,* 377-394.

Harris, G. T., Rice, M. E., & Quinsey, V. L. (1993). Violent recidivism of mentally disordered offenders. *Criminal Justice and Behavior, 20,* 315-335.

Harris, G. T., Skilling, T. A., & Rice, M. E. (2001). The construct of psychopathy. In M. Tonry & N. Morris (Eds.), *Crime and justice: An annual review of research* (pp. 197-264). Chicago: University of Chicago Press.

Heckert, D.A., & Gondolf, E.W. (2004). Battered women's perceptions of risk versus risk factors and instruments in predicting repeat reassault. *Journal of Interpersonal Violence, 19,* 778-800.

Hilton, N. Z., & Harris, G. T. (2005). Predicting wife assault: A critical review and implications for policy and practice. *Trauma, Violence, and Abuse, 6,* 3-23.

Hilton, N. Z., Harris, G. T., & Rice, M. E. (2001). Predicting violence by serious wife assaulters. *Journal of Interpersonal Violence, 16,* 408-423.

Hilton, N. Z., Harris, G. T., & Rice, M. E. (in press). The effect of arrest on wife assault recidivism, controlling for pre-arrest risk. *Criminal Justice and Behavior.*

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.

Hilton, N. Z., & Simmons, J. L. (2001). The influence of actuarial risk assessment in clinical

    judgments and tribunal decisions about mentally disordered offenders in maximum

    security. *Law and Human Behavior, 25,* 393-408.

Huss, M. T., Langhinrichsen-Rohling, J. (2000). Identification of the psychopathic batterer: The

    clinical, legal, and policy implications. *Aggression and Violent Behavior, 5,* 403-422.

Huss, M. T., Langhinrichsen-Rohling, J. (2006). Assessing the generalization of psychopathy in

    a clinical sample of domestic violence perpetrators. *Law and Human Behavior, 30,* 571-

    586.

Johnson, R., Gilchrist, E., Beech, A. R., Weston, S., Takriti, R., & Freeman, R. (2006). A

    psychometric typology of U.K. domestic violence offenders. *Journal of Interpersonal*

    Violence, 21, 1270-1285.

Kanji, G. K. (1993). *100 statistical tests.* London: Sage.

Koziol-McLain, J., Webster, D., McFarlane, J., Block, C. R., Ulrich, Y., Glass, N., et al. (2006).

    Risk factors for femicide-suicide in abusive relationships: Results from a multisite case

    control study. *Violence and Victims, 21,* 3-21.

Kropp, P. R., & Hart, S. D. (2000). The Spousal Assault Risk Assessment (SARA) Guide:

    Reliability and validity in adult male offenders. *Law and Human Behavior, 24,* 101-118.

Kropp, P. R., Hart, S. D., Webster, C. D., & Eaves, D. (1999). *Spousal Assault Risk Assessment*

    Guide. NY: Multi-Health Systems Inc.

Mossman, D. (2006). Another look at interpreting risk categories. *Sexual Abuse: Journal of*

    Research and Treatment, 18, 41-63.

Nunnally, J. C. (1959). *Tests and measurements: Assessment and prediction.* New York:

    McGraw-Hill.

Popham, S., & Hilton, N. Z. (2006, October). The *Ontario Domestic Assault Risk Assessment*

    (ODARA): Development and application to corrections. Paper presented at the biennial

    conference of the Ontario Ministry of Correctional Services and Community Safety,

    North Bay, ON.

Quinsey, V. L., Harris, G. T., Rice, M. E., & Cormier, C. A. (2006). *Violent offenders:*

    Appraising and managing risk. 2ⁿᵈ Ed. Washington, DC: American Psychological

    Association.

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.

Rice, M. E., Harris, G. T., Lang, C., & Cormier, C. A. (2006). Violent sex offenses: How are

    they best measured from official records? *Law and Human Behavior, 30,* 525-541.

Seto, M. (2005). Is more better? Combining actuarial risk scales to predict recidivism among

    adult sex offenders. *Psychological Assessment, 17*, 156-167.

Spidel, A., Vincent, G., Huss, M. T., Winters, J., Thomas, L., & Dutton, D. (2007). The

    psychopathic batterer: Subtyping perpetrators of domestic violence. In H. Herve and J. C.

    Yuille (Eds.) *The psychopath: Theory, research, and practice* (pp. 327-342). Mahwah,

    NJ: Lawrence Erlbaum.

Straus, M. A., Hamby, S. L., Boney-McCoy, S., & Sugarman, D. B. (1996). The Revised

    Conflict Tactics Scales (CTS2). *Journal of Family Issues, 17,* 283-316.

Weisz, A. N., Tolman, R. M., & Saunders, D. G. (2000). Assessing the risk of severe domestic

    violence. *Journal of Interpersonal Violence, 15,* 75-90.

Williams, K. R., & Grant, S. R. (2006). Empirically examining the risk of intimate partner

violence: The revised Domestic Violence Screening Instrument (DVSI-R). *Public Health

Reports, 131,* 400-408.

Williams, K. R., & Houghton, A. B. (2004). Assessing the risk of domestic violence reoffending:

A validation study. *Law and Human Behavior, 28,* 437-455.

Footnotes

[1]Because they are count variables and, therefore, skewed, some statisticians would advise

that the number of recidivistic incidents and the number of severe incidents not be subjected to

ordinary least-squares regression. To check against any questionable conclusions, we

transformed these outcome variables using a Poisson transformation before analyses. As a

second check, we subjected them to Poisson loglinear analyses in conjunction with ODARA

score and each of the candidate assessments in a main effects analysis. In all cases, the Poisson-

based analyses yielded the same results as the regression analyses reported here.

[2]For the weighted ODARA items, without PCL-R, ROC = .68 (SE = .03, 95% CI = .62 to

.74), d = .67, not a significant improvement over the unweighted ODARA.

[3]We examined many other available variables reflecting perpetrators' adult mental health,

early adjustment, childhood abuse and neglect, childhood exposure to domestic violence. The

available data did not permit the scoring of most psychiatric conditions (e.g., schizophrenia

symptoms, personality disorders). Other variables could be scored in at least 200 cases, but were

unrelated to wife assault recidivism (e.g., medical problems as an infant or young child;

experienced childhood corporal punishment, abuse, or neglect; witnessed domestic violence as a

child; prior head injury). Finally, a few variables scored for at least 200 cases were correlated

with dichotomous wife assault recidivism: as an adult, the perpetrator exhibited procriminal

attitudes and values, r (346) = .15, p < .01, and attitudes unfavorable to convention, r (335) =

.21, p <.01; had been suspended or expelled from elementary school, r (649) = .12, p < .01. All

had been previously identified as related to violent recidivism (Quinsey et al., 2006), but none

made an incremental improvement to DVRAG scores in predicting wife assault recidivism in

either Sample 1, Sample 2, or both samples combined. An exception, having been arrested under

age 16, r (649) = .10, $p < .05$, made a statistically significant improvement to DVRAG score in 3

of 15 tests – to the prediction of total victim injury in recidivism in the combined sample and in

Sample 2, and prediction of Cormier-Lang score for severity of recidivism in Sample 2.

Appendix: Scoring the Domestic Violence Risk Appraisal Guide (DVRAG)

Full scoring criteria are available from authors on request except where noted.

1. Number of prior domestic incidents (assault on a current or previous female cohabiting partner or her children, recorded in a police occurrence report or criminal record)

      0   = -1

      1   = 0

      ≥ 2 = +5

2. Number of prior nondomestic incidents (assault on any person other than a current or previous female cohabiting partner or her children, recorded in a police occurrence report or criminal record)

      0   = -1

      ≥ 1 = +5

3. Prior correctional sentence of 30 days or more

      No   = -1

      Yes = +2

4. Prior conditional release failure

      No   = -1

      Yes = +2

5. Threat to harm or kill at the index incident (threat of physical harm made towards any person other than himself)

      No   = 0

      Yes = +1

6. Confinement at the index incident (any attempt to physically prevent the female victim from leaving the scene of the incident)

    No = 0

    Yes = +1

7. Victim concern (concern, fear, worry, or certainty about possible future domestic assault, stated at the time of the index incident)

    No = 0

    Yes = +2

8. Number of children

    ≤ 1 = -1

    ≥ 2 = +1

9. Victim's number of biological children from a previous partner

    0 = -1

    1 = 0

    ≥ 2 = +2

10. Violence against others (any assault on any person other than a current or previous female cohabiting partner or her children)

    No = 0

    Yes = +8

11. Substance abuse score

    One point is allotted for each of the following: alcohol involved in the index incident, drugs involved in the index incident, alcohol or drug abuse in days/weeks prior to index incident, increased drug or alcohol use in days/weeks prior to index incident, more angry or violent

when using drugs or alcohol, alcohol involved in a prior criminal offense, adult alcohol

problem, adult drug problem.

≤ 1 = -2

≥ 2 = +2

12. Assault on victim when pregnant (index assault or prior)

No = 0

Yes = +5

13. Number of barriers to victim support

One point is allotted for each of the following: victim has children aged ≤ 18 to care for;

victim has no telephone or transportation; victim is isolated geographically or from

community; victim alcohol use in the index incident or victim adult alcohol or drug problem.

0 = -1

1 = 0

≥ 2 = +4

14. Psychopathy Checklist – Revised Score (full scoring criteria available in Hare, 2003)

≤ 9 = -1

10-16 = +1

≥ 17 = +6

Author Note

Table 1

Sample Characteristics, Shown as Mean (SD) or Percent of Sample

	Sample 1	Sample 2
Offender Characteristics		
Age (yr)	35.5 (10.1)	35.3 (10.0)
Unemployed (%)	14	14
Substance abuse score	1.94 (1.64)	1.60 (1.63)**
Number prior domestic incidents	0.16 (0.56)	0.27 (0.71)*
Prior criminal history score	9.95 (15.1)	16.3 (24.7)***
Violation of prior conditional release order (%)	42	27***
Violation of prior no-contact order (%)	8	4*
Total prior injury to female domestic partners	0.53 (1.50)	0.70(1.52)
Total prior injury to nondomestic victims	0.22 (1.23)	0.37(1.33)
Relationship Characteristics		
Victim age (yr)	32.4 (9.56)	32.4 (9.64)
Victim unemployed (%)	30.1	23.7
Duration of relationship (mo)	80.8 (83.9)	99.8 (106)*
Legally married at index (%)	37	45
Separated prior to index (%)	33	29
Perpetrator demonstrated jealousy (%)	14	11

Table continues

	Sample 1	Sample 2
Index Assault Details		
Weapon used (%)	6	8
Perpetrator charged (%)	75	90***
Injury to victim	2.08 (0.89)	2.19 (0.97)
Formal Assessments (potential range)		
ODARA (0 – 13)	4.05 (2.15)	3.54 (2.00)*
SARA (0 – 40)	4.63 (4.94)	4.01 (4.00)
DA (0 – 15)	0.73 (1.20)	0.75 (1.23)
DVSI (0 – 30)	2.68 (2.38)	2.25 (2.09)*
PCL-R (0 – 40)	8.00 (6.81)	8.35 (6.67)
VRAG (-26 – +38)	-2.17 (6.73)	-3.66 (6.56)*
Wife Assault Recidivism		
Any incident (%)	49	41
Number of incidents	0.83 (1.49)	0.65 (1.88)
Injury to partners (total score)	1.46 (2.20)	3.19 (3.40)***
Number of incidents with severe violence	0.30 (0.62)	0.70 (0.88)***
Cormier Lang score for recidivistic charges	1.46 (2.51)	1.69 (6.06)

Note: $*p < .05$, $**p < .01$, $***p < .001$. All characteristics except recidivism were coded as of date of index assault. Sample differences indicated by (t-test or Chi-square).

Table 2

Bivariate Correlations Between Each Candidate Formal Assessment and Each Outcome

Variable (nonsignificant correlations in italics) in Sample 1 (n = 303).

	VRAG	PCL-R	SARA	DA	DVSI
Dichotomous wife assault recidivism	.19***	.22***	.18**	.12*	.17**
Number of recidivistic incidents	.22***	.28***	.22***	*.07*	.24***
Total victim injury in recidivism	.28***	.31***	.21***	.17**	.18**
Number of CTS severe incidents	.22***	.26***	.20**	*.07*	.19**
Cormier-Lang score for recidivism	.14*	.23***	.18**	*.05*	.26***

Note: * $p < .05$, **$p < .01$, ***$p < .001$. Correlations are point-biserial for dichotomous

recidivism. For comparison, the ODARA, constructed partly on Sample 1, yielded correlations of

.27 to .34, ps < .001.

Table 3

Cross-validation Predictive Accuracy for the ODARA and DVRAG, and Z Scores of Difference

Between Correlations of ODARA vs DVRAG with Outcomes, in Sample 2 (n = 346).

	ODARA	DVRAG	*z* score
Dichotomous wife assault recidivism ROC area	.65	.70	--
Dichotomous wife assault recidivism	.29	.36	1.68
Number of recidivistic incidents	.36	.44	2.04
Total victim injury in recidivism	.34	.41	1.74
Number of CTS severe incidents	.26	.32	*1.46*
Cormier-Lang score for all recidivism	.24	.29	*1.15*

Note: All rows under ODARA and DVRAG columns except first row are Pearson correlation coefficients (point-biserial for dichotomous recidivism), all $p < .001$. Z scores $p < .05$ one-tailed except those in italics.

Table 4

Interpretation of DVRAG Scores in the Combined Samples (N = 649)

Score	Category	Cumulative Proportion	Overall Recidivism Rate	95%CI
-10 to -9	1	.02	.14	$\pm$.21
-8 to -5	2	.22	.24	$\pm$.07
-4 to -1	3	.43	.34	$\pm$.08
0 to +3	4	.63	.44	$\pm$.09
+4 to +10	5	.81	.51	$\pm$.09
+11 to +23	6	.97	.71	$\pm$.09
+24 to +41	7	100	1.0	0.0

Note: Observed DVRAG scores ranged from -10 to +37.

Table 5

Accuracy of Several Formal Risk Assessments (plus the PCL-R) in Predicting Wife Assault

Recidivism Outcome Variables in Samples 1 and 2 Combined (N = 649).

	VRAG	PCL-R	SARA	DA	DVSI
Dichotomous wife assault recidivism ROC area	.67	.66	.59	.56*	.61
Dichotomous wife assault recidivism	.29	.29	.21	.17	.20
Number of recidivistic incidents	.32	.36	.27	.24	.22
Total victim injury in recidivism	.31	.37	.22	.21	.17
Number of CTS severe incidents	.25	.29	.23	.22	.14
Cormier-Lang score for all recidivism	.23	.26	.24	.23	.20

Note: All rows except first row are Pearson correlation coefficients (point-biserial for dichotomous recidivism), $p < .01$ except where noted. * $p < .05$. For comparison, the ODARA and DVRAG, constructed partly on Sample 1, yielded ROC areas for dichotomous recidivism in the combined sample of .67 and .71, respectively; and correlations of .23 to .34, and .28 to .42, respectively, with the continuous outcome variables, all ps $< .001$.

Table 6

Inter-relationships (Pearson Correlation Coefficients) Among Formal Assessments in the

Combined Samples (N=649).

	ODARA	VRAG	PCL-R	SARA	DA	DVSI	DVRAG
ODARA	-	.53	.55	.60	.43	.52	.87
VRAG			.72	.43	.21	.31	.66
PCL-R				.55	.36	.34	.72
SARA					.61	.53	.63
DA						.36	.46
DVSI							.50
DVRAG							-

Note: All $p < .001$.