

# Allegiance or Fidelity? A Clarifying Reply

Grant T. Harris and Marnie E. Rice, Mental Health Centre, Penetanguishene  
Vernon L. Quinsey, Queen's University

**Recently, in the journal *Clinical Psychology: Science and Practice*, there appeared a systematic review (Blair, Marcus, & Boccaccini, 2008) accompanied by a commentary (Lilienfeld & Jones, 2008) suggesting an “allegiance effect” in the reporting of the predictive accuracy of actuarial risk assessment systems. The authors of these two articles suggested some possible errors or misrepresentation on the part of original developers or other researchers and proposed some remedies. We examined these two articles in conjunction with all the available evidence for the *Violence Risk Appraisal Guide* and *Sex Offender Risk Appraisal Guide* and concluded there is no evidence of an allegiance effect.**

**Key words:** meta-analysis, SORAG, violence risk assessment, VRAG. [*Clin Psychol Sci Prac* 17: 82–89, 2010]

Recently, in the pages of *Clinical Psychology: Science and Practice*, there appeared a meta-analysis accompanied by a commentary (Blair, Marcus, & Boccaccini, 2008; Lilienfeld & Jones, 2008) suggesting an “allegiance effect” in the reporting of empirical studies of actuarial risk assessments, such as those we developed (Quinsey, Harris, Rice, & Cormier, 2006). These instruments of ours, the *Violence Risk Appraisal Guide* (VRAG) and *Sex Offender Risk Appraisal Guide* (SORAG), clearly predicted relevant outcomes (Blair et al., 2008), and no superior method for assessing the risk of violence among forensic cases was proposed. However, the

---

Address correspondence to Grant T. Harris, MHCP, Research Department, 500 Church Street, Penetanguishene, ON L9M 1G3. E-mail: gharris@mhcp.on.ca.

authors of the meta-analysis reported that studies by the VRAG and SORAG designers yielded significantly larger predictive effects, mean  $r = .36$ , than those of other researchers, mean  $r = .30$ . They tested some possible explanations for this average difference but found empirical support for none.

The accompanying commentary (Lilienfeld & Jones, 2008), while reserving judgment as to whether this allegiance would be observed with other assessments, suggested but did not test other explanations for the findings of Blair et al. (2008) and for allegiance effects in general. Some of these reflected inadvertent error or misrepresentation by test designers or others—data “massaging,” selective reporting, other researchers’ allegiance to their own assessments. (Blair et al., 2008, also acknowledged a possible “allegiance-to-allegiance” on the part of meta-analysts.) Some general remedies were also described. The applicability of the commentators’ explanations and the value of their recommendations (whether for psychological assessment, in general, or actuarial tools, in particular) depend on there being allegiance effects in the first place. As our summaries have not indicated allegiance (Rice, 2008), we examined the issue further.

## **DIFFERENCES WITH BLAIR ET AL. (2008)**

Because statistics derived from Relative Operating Characteristics (ROCs) are preferred over correlation coefficients (Leistico, Salekin, deCoster, & Rogers, 2008; Rice & Harris, 1995, 2005), we analyzed areas under the ROC using standard conversions where required (Rice & Harris, 2005) instead of Pearson correlation coefficients<sup>1</sup> as used by Blair et al. (2008). We agree with the meta-analysts and commentators (Blair et al., 2008; Lilienfeld & Jones, 2008) that test

designers might well stick more closely to recommended scoring procedures and data sources than other researchers. Thus, we compared all studies by ourselves with those by others, and by categorizing others' studies as using scoring methods and outcomes similar to those recommended for the VRAG and SORAG (Quinsey, Harris, Rice, & Cormier, 1998, 2006), or very different methods and outcomes (Table 1).

#### **Different Outcomes**

Actuarial assessments are explicitly designed and recommended for specific outcomes—violent recidivism in the case of the VRAG and SORAG, and “sexual” recidivism (offenses that can be identified as contact or noncontact sexual crimes on police rap sheets) in the case of the Static-99. Blair et al. (2008) stated that “there is not always a clear reason for selecting one type of recidivism over another” (p. 350). Then, citing a meta-analysis (Leistico et al., 2008) of the Hare Psychopathy Checklist (PCL-R; Hare, 2003), they averaged over various outcomes within single studies. This decision struck us as inappropriate because the PCL-R was not designed as a structured risk assessment for any particular outcome while the converse is true of the VRAG and SORAG. They were both explicitly designed to predict violent (including violent sexual) criminal offenses after release to the community—the specific outcome is critically important. It is the prediction of that outcome that ought to have been examined if possible. Here, if a study reported on the prediction of violence in general, that was examined. If other outcomes were reported instead, we counted that which reflected some kind of violence. In such cases, the study was assigned to the “Very Different Procedures and/or Outcomes” category. Some studies in that category examined general criminal recidivism (Bélanger & Earls, 1996), only “sexual” recidivism (whether violent or not; e.g., Hanson & Harris, 2000), or the first detected recidivism whether or not violence occurred afterwards (Loza, Villeneuve, & Loza-Fanous, 2002). We also counted in the “Very Different Outcomes” a study of sex offenders that used nonsexual violent recidivism as the outcome (Sjostedt & Langstrom, 2002). Five studies that employed only institutional misconduct as the outcome formed a fourth category (Table 1).

#### **Different Procedures**

Two studies in the “Very Different Procedures and/or Outcomes” category used highly nonstandard scoring procedures. That is, Tengstrom (2001) scored two items as constants, three were approximated, and one was apparently reverse scored. Doyle and Dolan (2006) based scoring and outcome on interviews and gave the lowest possible score for three missing items (instead of zero or prorating as set out in instructions; Quinsey et al., 1998, 2006).

#### **The VRAG and SORAG as Different**

Blair et al. (2008) treated findings separately for the VRAG and SORAG when both were reported in a single study. The VRAG and SORAG share most of their items and are very highly correlated (approximately 0.9; Harris et al., 2003; Langton et al., 2007; Yaghoub, Fedoroff, Curry, & Amundsen, in press). Thus, we averaged predictive effects across the two whenever they were evaluated in one study. This always involved differences  $< 0.05$  (in ROC area).

#### **Missed Studies, New Studies, and Overlap**

Blair et al. (2008) referred to a website that has listed all known tests pertaining to the accuracy of the VRAG and SORAG (<http://www.mhcop-research.com/ragreprs.htm>). Nevertheless, their analysis missed several studies listed therein that met their inclusion criteria (e.g., Doyle & Dolan, 2006; Pham, Ducro, Marghem, & Réveillère, 2005; Quinsey, Book, & Skilling, 2004; Quinsey, Coleman, Jones, & Altrows, 1997; Quinsey, Jones, Book, & Barr, 2006). Several studies, mostly by others, have appeared since (e.g., Coid et al., 2009; Gray, Fitzgerald, Taylor, MacCulloch, & Snowden, 2007; Hilton, Harris, Rice, Houghton, & Eke, 2008; Kingston, Yates, Firestone, Babchishin, & Bradford, 2008; Kroner, Stadtland, Eidt, & Nedopil, 2007; Rettenberger & Eher, 2007; Rice, Harris, Lang, & Chaplin, 2008; Snowden, Gray, Taylor, & MacCulloch, 2007). Blair et al. (2008) considered Quinsey, Rice, and Harris (1995) as a study of the SORAG; it was not. Also, Hanson and Harris (2000) was not by us—it is A. J. R. Harris (no relation).

In addition, we examined a published, peer-reviewed study in preference to an unpublished report (Yessine & Bonta, 2006, versus Bonta & Yessine,

**Table 1.** All known tests of the *Violence Risk Appraisal Guide* (VRAG) and *Sex Offender Risk Appraisal Guide* (SORAG)

|   | N      | ROC  | Comments (v) = VRAG, (s) = SORAG, (vs) = both, averaged  |
|---|--------|------|--|
| <b>Group 1: Developers</b>                                |        |      |  |
| Rice & Harris, 1995                                       | 799    | 0.74 | Includes development samples (v)   |
| Quinsey et al., 1997                                      | 69     | 0.68 | Matching on age and diagnosis; not a follow-up (v)   |
| Rice & Harris, 1997                                       | 159    | 0.77 | Sex offenders, but only those not in Rice & Harris, 1995 (v)                                     |
| Glover et al., 2002                                       | 106    | 0.72 | (v)  |
| Harris et al., 2002                                       | 347    | 0.75 | Only males; higher ROC with constant follow-up but not listed here; hospital diagnosis (v)       |
| Harris et al., 2003                                       | 396    | 0.73 | Sex offenders from three institutions (vs)   |
| Quinsey et al., 2004                                      | 58     | 0.69 | PCL-R replaced with CATS (v)   |
| Harris et al., 2004                                       | 741    | 0.72 | PCL-SV; missing 2 VRAG items; others approximated; mostly self-report violent recidivism (v)     |
| Quinsey, Jones, et al., 2006                              | 198    | 0.59 | Used VRAG categories (v)   |
| Hilton et al., 2008                                       | 649    | 0.67 | Missing three VRAG items; domestic violence recidivism (v)                                       |
| Rice et al., 2008   | 61     | 0.67 | SORAG not available for several participants (v)   |
| Total   | 3,583  | 0.71 | Total N and weighted mean ROC area   |
| <b>Group 2: Similar methods and violent recidivism</b>    |        |      |  |
| Dempster, 1998  | 95     | 0.85 | Unpublished master's thesis; used VRAG and SORAG categories (vs)                                 |
| Polvi, 2001   | 215    | 0.70 | Unpublished dissertation (v)   |
| Cooke et al., 2001  | 250    | 0.71 | Not peer reviewed (v)  |
| Nugent, 2001  | 123    | 0.68 | Unpublished dissertation (v)   |
| Pham, 2002  | 58     | 0.84 | Unpublished conference presentation (v)  |
| Bartosh et al., 2003                                      | 186    | 0.64 | PCL-R replaced with CATS; missing two other SORAG items (s)                                      |
| Mills et al., 2005  | 209    | 0.65 | Volunteer subjects (v)   |
| Douglas et al., 2005                                      | 188    | 0.79 | Not a follow-up (v)  |
| Pham et al., 2005   | 114    | 0.82 | (v)  |
| Ducro & Pham, 2006  | 133    | 0.72 | (s)  |
| Looman, 2006  | 242    | 0.70 | Article has an error in computation of ROC area (s)  |
| Hastings et al., 2006                                     | 326    | 0.74 | Self-reported violent recidivism (v)   |
| Urbanik et al., 2006                                      | 79     | 0.72 | (v)  |
| Yessine & Bonta, 2006                                     | 165    | 0.73 | PCL-R missing for most cases; other VRAG items approximated (v)                                  |
| Johansen, 2007  | 280    | 0.76 | PCL-R replaced with CATS; unpublished dissertation (vs)  |
| Kroner et al., 2007                                       | 113    | 0.70 | (v)  |
| Langton et al., 2007                                      | 468    | 0.70 | Mostly treated sex offenders (vs)  |
| Snowden et al., 2007                                      | 421    | 0.76 | (v)  |
| Gray et al., 2007   | 404    | 0.74 | Approximately 20% female (v)   |
| Rettenberger & Eher, 2007                                 | 254    | 0.76 | (s)  |
| Thomson et al., 2008                                      | 135    | 0.80 | Forensic patients with schizophrenia; release positively associated with VRAG scores (v)         |
| Rosales & Rossegger, 2008                                 | 107    | 0.62 | Conference abstract (v)  |
| Kingston et al., 2008                                     | 192    | 0.76 | Modifications to some SORAG items (s)  |
| Coid et al., 2009   | 1,645  | 0.69 | Scoring based on interview (v)   |
| Ho et al., 2009   | 96     | 0.74 | Several other outcomes yielded smaller effects (v)   |
| Rettenberger et al., in press                             | 394    | 0.72 | (s)  |
| Yaghoub et al., in press                                  | 526    | 0.72 | Unusual definition of static and dynamic factors (vs)  |
| Total   | 7,418  | 0.72 | Total N and weighted mean ROC area   |
| <b>Group 3: Very different procedures and/or outcomes</b> |        |      |  |
| Bélanger & Earls, 1996                                    | 57     | 0.82 | General recidivism; violent recidivism not reported; not peer reviewed (v)                       |
| Hanson & Harris, 2000                                     | 267    | 0.70 | Not a follow-up; rap sheet sexual recidivism; violent recidivism not reported (v)                |
| Grann et al., 2000  | 404    | 0.68 | At least 3 VRAG items missing or approximated; violent recidivism inconsistently defined (v)     |
| Hartwell, 2001  | 164    | 0.67 | Rap sheet sexual recidivism; violent recidivism not reported; unpublished manuscript (s)         |
| Tengstrom, 2001   | 106    | 0.68 | Several VRAG items missing/approximated or reversed (v)  |
| Sjostedt & Langstrom, 2002                                | 51     | 0.69 | Several VRAG items missing/approximated; low reliability; nonsexual violent recidivism (v)       |
| Loza et al., 2002   | 124    | 0.54 | Volunteers; by interview; counted first recorded recidivism regardless of later violence (v)     |
| Doyle & Dolan, 2006                                       | 112    | 0.66 | For three missing VRAG items used lowest possible score; interview-based outcome (v)             |
| Knight & Thornton, 2007                                   | 537    | 0.62 | Rap sheet sexual recidivism; violent recidivism not reported; not peer reviewed (vs)             |
| Total   | 1,822  | 0.66 | Total N and weighted mean ROC area   |
| <b>Group 4: Institutional misconduct</b>                  |        |      |  |
| Doyle et al., 2002  | 87     | 0.71 | Included threatened and attempted violence (v)   |
| Endrass et al., 2008                                      | 106    | 0.50 | Estimated; all they say is that it was not related for violence (v)                              |
| McDermott et al., 2008                                    | 108    | 0.54 | Volunteers, some female, staff and patient victims combined (v)                                  |
| Lindsay et al., 2008                                      | 212    | 0.71 | Violence scored from nursing notes (v)   |
| Snowden et al., 2009                                      | 52     | 0.77 | Violence scored from nursing notes; also a large effect predicting frequency of violent acts (v) |
| Total   | 565    | 0.64 | Total N and weighted mean ROC area   |
| Overall   | 13,388 | 0.71 | Grand Total N; weighted mean ROC area; all methods, outcomes, and sources (k = 52)               |
| <b>Unclassifiable</b>                                     |        |      |  |
| McBride, 1999   | ?      | ?    | Includes institutional misconduct; unpublished conference presentation (v)                       |
| Nichols et al., 1999                                      | ?      | ?    | Institutional misconduct; unpublished conference presentation (v)                                |
| Nadeau et al., 1999                                       | ?      | 0.65 | Institutional misconduct; unpublished conference presentation (v)                                |
| Douglas et al., 1999                                      | ?      | 0.60 | Most VRAG items approximated; unpublished conference presentation (v)                            |

2005). Relatedly, as Blair et al. and Lilienfeld and Jones suggested the “file drawer” problem as a general explanation for possible allegiance effects, we could not agree with the decision to exclude unpublished findings. We examined all studies we could find, published or not.

Finally, the literature contains overlapping and extended sets of data (as acknowledged by Blair et al., 2008 cf. Barbaree, Seto, Langton, & Peacock, 2001, versus Langton et al., 2007). For example, the original construction sample for the VRAG (Harris, Rice, & Quinsey, 1993) was effectively superseded by an extended sample with a longer follow-up (Rice & Harris, 1995) so that we do not here consider the former separately. Similarly, the entire sample of sex offenders (Rice & Harris, 1997) overlapped with the latter sample (Rice & Harris, 1995). For present purposes (and in our reports about replications; Quinsey, Harris, et al., 2006), we examined only the nonoverlapping sub-sample as a replication of the VRAG’s predictive accuracy. In the present examination (as Blair et al. (2008) also attempted to do), we avoided such “double counting” by considering only the most complete among any overlapping samples (e.g., Loza & Loza-Fanous, 2001, plus Kroner & Loza, 2001, were superseded by Loza et al., 2002; Kroner & Mills, 2001, was superseded by Mills, Jones, & Kroner, 2005; Nunes, Firestone, Bradford, Greenberg, & Broom, 2002, was superseded by Kingston et al., 2008; and Rice & Harris, 2002, was superseded by Harris et al., 2003).

## RESULTS AND DISCUSSION

The results of our evaluation are shown in Table 1. Four studies we were aware of could not be classified (Douglas, Hart, Dempster, & Lyon, 1999; McBride, 1999; Nadeau, Nadeau, Smiley, & McHattie, 1999; Nichols, Vincent, Whittemore, & Ogloff, 1999); all were unpublished conference presentations. Each of the four classifiable groups of studies shown (total  $k = 52$ ) exhibited significant heterogeneity in effect size, chi-square statistics = 26.55 (d.f. = 10), 58.57 (d.f. = 26), 18.20 (d.f. = 8), and 19.14 (d.f. = 4), respectively, all  $p < .05$ .

We see no allegiance effect for violent recidivism. Readers can draw their own conclusions and might disagree with our unblinded classification of the 36 recidi-

vism studies by other researchers (27 in Group 2 and 9 in Group 3). However, averaging across these two groups yielded a mean weighted ROC area = 0.71 (compared to 0.71 for all 11 studies by developers, a few of which also used nonstandard scoring or outcomes other than violent recidivism in general). The weighted average of all 41 studies by others (in Groups 2, 3, and 4), irrespective of methods and outcomes, was 0.70. Regarding published, peer-reviewed studies versus unpublished findings, there were 41 instances of the former, weighted mean ROC area = 0.71, and 11 of the latter, weighted mean ROC area = 0.70. Nothing can be attributed to file drawers, especially as there was no allegiance effect. Overall, the VRAG and SORAG scored using recommended procedures for the assessment of the risk of violent community recidivism yielded a mean weighted ROC area greater than 0.72, a large effect by common standards in the behavioral sciences (Cohen, 1988; Rice & Harris, 2005).

The prediction by the VRAG and SORAG of violent community recidivism in the present systematic review is consistent with, indeed essentially identical to, that reported in meta-analyses by others (Hanson & Morton-Bourgon, 2009). The present review does make two things evident. First, although we occasionally deviate from our recommended procedures for scoring the VRAG and SORAG from comprehensive forensic files (e.g., Harris, Rice, & Camilleri, 2004), such deviations are larger and more frequent in studies conducted by others (e.g., Doyle & Dolan, 2006; Tengstrom, 2001). Second, we have tested the VRAG and SORAG principally for the purpose for which we designed them—criminally violent community recidivism. Although we have sometimes reported the prediction of other outcomes as well (e.g., Harris et al., 2003; Hilton et al., 2008), other researchers (e.g., Doyle, Dolan, & McGovern, 2002; Hanson & Harris, 2000; Knight & Thornton, 2007; McDermott, Quabeck, Scott, Edens, & Busse, 2008; Sjostedt & Langstrom, 2002) are more likely to have studied only outcomes the instruments were not designed to predict. Although it is laudable to study various outcomes, the unsystematic inclusion by Blair et al. (2008) of studies of various outcomes in evaluating the instruments’ predictive accuracy confounded study origin with the outcome for which the systems were explicitly designed.

We suggest these differences between our own research and that of a few others are a sort of “allegiance,” but it is fidelity to validated procedures (a possibility acknowledged by Lilienfeld & Jones, 2008) and intended uses as set out in our published work (Harris et al., 1993; Quinsey et al., 1998, 2006), rather than the errors or misrepresentations they also suggested as the basis for allegiance effects. Readers may, of course, use Table 1 to examine other sources of variation in predictive effects. Also, as illustrated by Blair et al. (2008), sorting out study origins and sample overlap can be tricky. In the end, however, we conclude that examining all currently available data yields no evidence of an allegiance effect, and therefore, no basis in our work for a commentary (Lilienfeld & Jones, 2008) on the topic.

#### NOTE

1. Blair et al. (2008) neglected the fact that the correlations they reported were point-biserial  $r$  because outcomes were dichotomous. They did not convert or adjust their correlation coefficients between studies to compensate for the dependency between such coefficients and individual study base rates (Rice & Harris, 2005). The present analyses of heterogeneity used standard methods based on Cohen’s  $d$  statistic (Wolf, 1986).

#### REFERENCES

- Barbaree, H. E., Seto, M. C., Langton, C. M., & Peacock, E. J. (2001). Evaluating the predictive accuracy of six risk assessment instruments for adult sex offenders. *Criminal Justice and Behavior*, 28, 490–521.
- Bartosh, D. L., Garby, T., Lewis, D., & Gray, S. (2003). Differences in the predictive validity of actuarial risk assessments in relation to sex offender type. *International Journal of Offender Therapy and Comparative Criminology*, 47, 422–438.
- Bélanger, N., & Earls, C. (1996). Sex offender recidivism prediction. *Forum on Correctional Research*, 8, 22–24.
- Blair, P. R., Marcus, D. K., & Boccaccini, M. T. (2008). Is there an allegiance effect for assessment instruments? Actuarial risk assessment as an exemplar. *Clinical Psychology: Science and Practice*, 15, 346–360.
- Bonta, J., & Yessine, A. K. (2005). *The national flagging system: Identifying and responding to high risk, violent offenders*. (Cat. No.: PS3-1/2005-3E-PDF). Ottawa, Canada: Public Works and Government Services Canada.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.
- Coid, J., Yang, M., Ullrich, S., Sizmur, S., Roberts, C., Farrington, D. P., et al. (2009). Gender differences in structured risk assessment: Comparing the accuracy of five instruments. *Journal of Consulting and Clinical Psychology*, 77, 337–348.
- Cooke, D. J., Michie, C., & Ryan, J. (2001). *Evaluating risk for violence: A preliminary study of the HCR-20, PCL-R and VRAG in a Scottish prison sample*. Scottish Prison Service Paper; no. 5/2001. ISBN 0954101022
- Dempster, R. J. (1998). *Prediction of sexually violent recidivism: A comparison of risk assessment instruments*, Unpublished master’s thesis, Simon Fraser University, Burnaby, British Columbia, Canada.
- Douglas, K. S., Hart, S. D., Dempster, R. J., & Lyon, D. R. (1999). Violence Risk Appraisal Guide (VRAG): Attempt at validation in a maximum-security forensic psychiatric sample. Paper presented at the joint meeting of the American Psychology-Law Society and the European Association of Psychology and Law, Dublin, Eire.
- Douglas, K., Yeomans, M., & Boer, D. F. (2005). Comparative validity analysis of multiple measures of violence risk in a sample of criminal offenders. *Criminal Justice and Behavior*, 32, 479–510.
- Doyle, M., & Dolan, M. (2006). Predicting community violence from patients discharged from community mental health services. *British Journal of Psychiatry*, 189, 520–526.
- Doyle, M., Dolan, M., & McGovern, J. (2002). The validity of North American risk assessment tools in predicting in-patient violent behaviour in England. *Legal & Criminological Psychology*, 7, 141–152.
- Ducro, C., & Pham, T. (2006). Evaluation of the SORAG and the Static-99 on Belgian sex offenders committed to a forensic facility. *Sexual Abuse: A Journal of Research and Treatment*, 18, 15–25.
- Endrass, J., Rossegger, A., Frischknecht, A., Noll, T., & Urbaniok, F. (2008). Using the Violence Risk Appraisal Guide (VRAG) to predict in-prison aggressive behavior in a Swiss offender population. *International Journal of Offender Therapy Comparative Criminology*, 52, 81–89.
- 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.
- Grann, M., Belfrage, H., & Tengstrom, A. (2000). Actuarial assessment of risk for violence: Predictive validity of the VRAG and historical part of the HCR-20. *Criminal Justice and Behavior*, 27, 97–114.
- Gray, N. S., Fitzgerald, S., Taylor, J., MacCulloch, M. J., & Snowden, R. J. (2007). Predicting future reconviction in



- offenders with intellectual disabilities: The predictive efficacy of VRAG, PCL-SV and the HCR-20. *Psychological Assessment*, 19, 474–479.
- Hanson, R. K., & Harris, A. J. R. (2000). Where should we intervene? Dynamic predictors of sex offense recidivism. *Criminal Justice and Behavior*, 27, 6–35.
- Hanson, R. K., & Morton-Bourgon, K. E. (2009). The accuracy of recidivism risk assessments for sexual offenders: A meta-analysis. *Psychological Assessment*, 21, 1–21.
- Hare, R. D. (2003). *The Psychopathy Checklist—Revised manual* (2nd ed.). Toronto, Canada: Multi-Health Systems.
- 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: The development of a statistical prediction instrument. *Criminal Justice and Behavior*, 20, 315–335.
- Harris, G. T., Rice, M. E., Quinsey, V. L., Lalumière, M. L., Boer, D., & Lang, C. (2003). A multi-site comparison of actuarial risk instruments for sex offenders. *Psychological Assessment*, 15, 413–425.
- Hartwell, L. L. (2001). *Sex Offender Risk Appraisal Guide: Validity and utility for Hawaii sex offender risk assessments*. Unpublished clinical research project, American School of Professional Psychology, Hawaii Campus.
- Hastings, M. E., Stuewig, J., & Tangney, J. (2006). *Relative efficacy of the PAI, PCL: SV, and VRAG in predicting institutional misconduct and short-term recidivism*. Paper presented at the annual meeting of the American Psychology-Law Society, St. Petersburg, FL.
- Hilton, N. Z., Harris, G. T., Rice, M. E., Houghton, R. E., & Eke, A. W. (2008). An indepth actuarial assessment for wife assault recidivism: The Domestic Violence Risk Appraisal Guide. *Law and Human Behavior*, 32, 150–163.
- Ho, H., Thomson, L., & Darjee, R. (2009). Violence risk assessment: The use of the PCL-SV, HCR-20, and VRAG to predict violence in mentally disordered offenders discharged from a medium secure unit in Scotland. *The Journal of Forensic Psychiatry and Psychology*, 20, 523–541.
- Johansen, S. H. (2007). Accuracy of predictions of sexual offense recidivism: A comparison of actuarial and clinical methods. *Dissertation Abstracts International: Section B: The Sciences and Engineering*, 68 (3-B), 1929.
- Kingston, D. A., Yates, P. M., Firestone, P., Babchishin, K., & Bradford, J. M. (2008). Long-term predictive validity of the Risk Matrix 2000: A comparison with the Static-99 and the Sex Offender Risk Appraisal Guide. *Sexual Abuse: A Journal of Research and Treatment*, 20, 466–484.
- Knight, R. A., & Thornton, D. (2007). *Evaluating and improving risk assessment schemes for sexual recidivism: A long-term follow-up of convicted sexual offenders*. Unpublished report to the U.S. Department of Justice, Washington, DC.
- Kroner, D. G., & Loza, W. (2001). Evidence for the efficacy of self-report in predicting violent and nonviolent criminal recidivism. *Journal of Interpersonal Violence*, 16, 168–177.
- Kroner, D., & Mills, J. (2001). The relative efficacy of predicting criminal behavior: A comparison of five instruments. *Criminal Justice and Behavior*, 28, 471–489.
- Kroner, C., Stadtland, C., Eidt, M., & Nedopil, N. (2007). The validity of the Violence Risk Appraisal Guide (VRAG) in predicting criminal recidivism. *Criminal Behaviour and Mental Health*, 17, 89–100.
- Langton, C. M., Barbaree, H. E., Seto, M. C., Peacock, E. J., Harkins, L., & Hansen, K. T. (2007). Actuarial assessment of risk for reoffense among adult sex offenders. *Criminal Justice and Behavior*, 34, 37–59.
- Leistico, A. R., Salekin, R. T., deCoster, J., & Rogers, R. (2008). A large-scale meta-analysis relating the Hare measures of psychopathy to antisocial conduct. *Law and Human Behavior*, 32, 28–45.
- Lilienfeld, S., & Jones, M. K. (2008). Allegiance effects in assessment: Unresolved questions, potential explanations, and constructive remedies. *Clinical Psychology: Science and Practice*, 15, 361–365.
- Lindsay, W. R., Hogue, T., Taylor, J. L., Steptoe, L., Mooney, P., O'Brien, G., et al. (2008). Risk assessment in offenders with intellectual disability: A comparison across three levels of security. *International Journal of Offender Therapy and Comparative Criminology*, 52, 90–111.
- Looman, J. (2006). Comparison of two risk assessment instruments for sexual offenders. *Sexual Abuse: A Journal of Research and Treatment*, 18, 193–206.
- Loza, W., & Loza-Fanous, A. (2001). The effectiveness of the self-appraisal questionnaire in predicting offenders' postrelease outcome. *Criminal Justice and Behavior*, 28, 105–121.
- Loza, W., Villeneuve, D. B., & Loza-Fanous, A. (2002). Predictive validity of the Violence Risk Appraisal Guide: A tool for assessing violent offenders' recidivism. *International Journal of Law and Psychiatry*, 25, 85–92.

- McBride, M. (1999). *Predicting violence among federal inmates*. Toronto: Corrections Research Forum.
- McDermott, B. E., Quanbeck, C. D., Scott, C. L., Edens, F., & Busse, D. (2008). Examining the role of static and dynamic risk factors in the prediction of in-patient violence. *Law and Human Behavior, 32*, 325–338.
- Mills, J. F., Jones, M. N., & Kroner, D. G. (2005). An examination of the generalizability of the LSI-R and VRAG probability bins. *Criminal Justice and Behavior, 32*, 565–585.
- Nadeau, J., Nadeau, B., Smiley, W. C., & McHattie, L. (1999). *The PCL-R and VRAG as predictors of institutional behaviour*. Paper presented at the conference on “Risk assessment & risk management: Implications for the prevention of violence,” Vancouver, British Columbia.
- Nichols, T. L., Vincent, G. M., Whittemore, K. E., & Ogloff, J. R. P. (1999). *Assessing risk of inpatient violence in a sample of forensic psychiatric patients: Comparing the PCL:SV, HCR-20, and VRAG*. Paper presented at the conference on “Risk assessment and risk management: Implications for the prevention of violence,” Vancouver, British Columbia.
- Nugent, P. M. (2001). The use of detention legislation: Factors affecting detention decisions and recidivism among high-risk federal offenders in Ontario (Doctoral dissertation, Queen’s University at Kingston, Ontario, 2001). *Dissertation Abstracts International: Section B: The Sciences and Engineering, 61* (12-B), 6716.
- Nunes, K. L., Firestone, P., Bradford, J. M., Greenberg, D. M., & Broom, I. (2002). A comparison of modified versions of the Static-99 and the Sex Offender Risk Appraisal Guide. *Sexual Abuse: A Journal of Research and Treatment, 14*, 253–269.
- Pham, T. H. (2002). *Risk assessment and biological markers*. Paper presented at Belgian Congress of Neuropsychology.
- Pham, T. H., Ducro, C., Marghem, B., & Réveillère, C. (2005). Evaluation du risque de récidivie au sein d’une population de délinquants incarcéré ou interné en Belgique francophone. *Annales Médico Psychologiques, 163*, 842–845.
- Polvi, N. H. (2001). *The relative efficacy of statistical versus clinical predictions of dangerousness (Doctoral dissertation, Simon Fraser University, 2001)*. *Dissertation Abstracts International: Section B: The Sciences and Engineering, 61* (7-B).
- Quinsey, V. L., Book, A., & Skilling, T. A. (2004). A follow-up of deinstitutionalized men with intellectual disabilities and histories of antisocial behaviour. *Journal of Applied Research in Intellectual Disabilities, 17*, 243–253.
- Quinsey, V. L., Coleman, G., Jones, B., & Altrows, I. F. (1997). Proximal antecedents of eloping and reoffending among mentally disordered offenders. *Journal of Interpersonal Violence, 12*, 794–813.
- Quinsey, V. L., Harris, G. T., Rice, M. E., & Cormier, C. A. (1998). *Violent offenders: Appraising and managing risk*. Washington, DC: American Psychological Association.
- 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.
- Quinsey, V. L., Jones, B., Book, A. S., & Barr, K. N. (2006). The dynamic prediction of antisocial behavior among forensic psychiatric patients. *Journal of Interpersonal Violence, 21*, 1539–1565.
- Quinsey, V. L., Rice, M. E., & Harris, G. T. (1995). The actuarial prediction of sexual recidivism. *Journal of Interpersonal Violence, 10*, 85–105.
- Rettenberger, M., & Eher, R. (2007). Predicting reoffence in sexual offender subtypes: A prospective validation study of the German version of the Sex Offender Risk Appraisal Guide (SORAG). *Sexual Offender Treatment, 2*, 1–12.
- Rettenberger, M., Matthes, A., Boer, D., & Eher, R. (in press). Prospective actuarial risk assessment: A comparison of five risk assessment instruments in different sexual offender subtypes. *International Journal of Offender Therapy and Comparative Criminology*.
- Rice, M. E. (2008). Current status of violence risk assessment: Is there a role for clinical judgment? In G. Bourgon, R. K. Hanson, J. D. Pozzulo, K. E. Morton Bourgon, & C. L. Tanasichuk (Eds.), *Proceedings of the North American Correctional and Criminal Justice Psychology Conference* (pp. 20–23). Public Safety Canada User Report 2008-02. Ottawa, Ontario: Public Safety Canada.
- Rice, M. E., & Harris, G. T. (1995). Violent recidivism: Assessing predictive validity. *Journal of Consulting and Clinical Psychology, 63*, 737–748.
- Rice, M. E., & Harris, G. T. (1997). Cross validation and extension of the *Violence Risk Appraisal Guide* for child molesters and rapists. *Law and Human Behavior, 21*, 231–241.
- Rice, M. E., & Harris, G. T. (2002). Men who molest their sexually immature daughters: Is a special explanation required? *Journal of Abnormal Psychology, 111*, 329–339.
- 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., & Chaplin, T. C. (2008). The sexual preferences and recidivism of sex offenders with mental retardation. *Sexual Abuse: A Journal of Research and Treatment, 20*, 409–425.

- Rosales, A. V., & Rossegger, A. (2008). *Validation and calibration of the VRAG in Switzerland*. Presentation to the European Congress of Psychiatry. Nice, France.
- Sjostedt, G., & Langstrom, N. (2002). Assessment of risk for criminal recidivism among rapists: A comparison of four different measures. *Psychology, Crime and Law, 8*, 25–40.
- Snowden, R. J., Gray, N. S., Taylor, J., & Fitzgerald, S. (2009). Assessing risk of future violence among forensic psychiatric inpatients with the Classification of Violence Risk (COVR). *Psychiatric Services, 60*, 1523–1526.
- Snowden, R. J., Gray, N., Taylor, J., & MacCulloch, M. J. (2007). Actuarial prediction of violent recidivism in mentally disordered offenders. *Psychological Medicine, 37*, 1539–1549.
- Tengstrom, A. (2001). Long-term predictive validity of historical factors in two risk assessment instruments in a group of violent offenders with schizophrenia. *Nordic Journal of Psychiatry, 55*, 243–249.
- Thomson, L., Davidson, M., Brett, C., Steele, J., & Darjee, R. (2008). Risk assessment in forensic patients with schizophrenia: The predictive validity of actuarial scales and symptom severity for offending and violence over 8–10 years. *International Journal of Forensic Mental Health, 7*, 173–189.
- Urbaniok, F., Noll, T., Grunewald, S., Steinbach, J., & Endrass, J. (2006). Prediction of violent and sexual offences: A replication study of the VRAG in Switzerland. *Journal of Forensic Psychiatry and Psychology, 17*, 23–31.
- Wolf, F. M. (1986). *Meta-analysis: Quantitative methods for research synthesis*. Newbury Park, CA: Sage.
- Yaghoub, M. B., Fedoroff, J., Curry, S., & Amundsen, D. E. (in press). A time series modeling approach in risk appraisal of violent and sexual recidivism. *Law and Human Behavior*.
- Yessine, A. K., & Bonta, J. (2006). Tracking high-risk, violent offenders: An examination of the national flagging system. *Canadian Journal of Criminology and Criminal Justice, 48*, 573–607.

Received November 9, 2009; accepted November 12, 2009.