Risk relevance of psychometric assessment and evaluator ratings of dynamic risk factors in high-risk violent offenders

Tamsin Higgs1,* Mark E. Olver2, Kevin Nunes3 and Franca Cortoni1

1University of Montreal, Quebec, Canada
2University of Saskatchewan, Saskatoon, Canada
3Carleton University, Ottawa, Ontario, Canada

Purpose. Relatively little research has been conducted with high-risk violent (non-sexual) offenders to establish whether measures administered to evaluate change during offending behaviour programmes contain risk relevant information. The present study aims to contribute to the evidence base relevant to decisions concerning whether or not psychometric assessments indicate how the violence risk presented by an individual may be understood differently pre- to post-treatment.

Methods. Two hundred and twenty-seven persistently violent offenders participating in Correctional Service of Canada’s Violence Prevention Program were assessed on measures of anger, impulsivity, and dynamic items of the Violence Risk Scale (VRS; Wong & Gordon, 1999–2003; Violence Risk Scale, University of Saskatchewan, Saskatchewan, CA) prior to and after programme completion and subsequently followed up in the community for an average of 3 years. Data were examined using receiver operating characteristic and logistic regression analyses employing fixed follow-ups.

Results. With pre-treatment status controlled, change on few of the measures convincingly predicted violent or general recidivism. An exception was that changes in VRS dynamic score were associated with decreased general (but not violent) recidivism, controlling for baseline pre-treatment risk.

Conclusions. The measures tested are widely used to evaluate progress in violence interventions yet the implicit assumption that they contain risk relevant information has not been empirically validated. Since reduction in dynamic risk factors translates into reduced likelihood of reoffending, but psychometric measures provide little indication of change in recidivism risk, treatment providers are advised to carefully contextualize pre- to post-treatment change within a comprehensive evaluation of static and dynamic risk using a measure such as the VRS. Present results are discussed further in terms of implications for policy and clinical practice, as well as future research directions.

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

*Correspondence should be addressed to Tamsin Higgs, University of Montreal, Pavillon Marie-Victorin, 90, avenue Vincent d’Indy, Montréal, QC H2V 2S9, Canada (e-mail: tamsin.higgs@umontreal.ca).

DOI:10.1111/lcrp.12173
There is now a fairly sizeable literature tending to demonstrate favourable outcomes for offenders who participate in offending behaviour programmes compared to those who do not (Andrews & Bonta, 2010a; Hanson, Bourgon, Helmus, & Hodgson, 2009; Henwood, Chou, & Browne, 2015; Lösel & Schmucker, 2005). While a general conclusion that ‘treatment works’ requires scrutiny given the methodological limitations inherent in research in the domain (Nunes et al., 2019), there remains a need to further elucidate precisely what works, with whom, and how (McGuire, 2001; Palmer, 1975; Ward, 2016). The effectiveness of interventions aiming to reduce risk of reoffending is sensitive to the duration and intensity of treatment according to criminal propensity, degree, and complexity of psychosocial difficulties associated with offending behaviour, and the appropriateness of interventions depending on systemic and individual factors. Greater confidence may be had in programmes that adhere to these risk, need, and responsivity principles (Andrews & Bonta, 2010b; Hanson et al., 2009).

Increasingly, as correctional programmes develop in line with empirical advances, programmes will successfully target dynamic factors, which, if truly dynamic and risk relevant, will mean reduced risk of reoffending when change is achieved in a sustainable way. The veritable dynamic nature of risk factors for offending behaviour has been largely assumed rather than empirically tested (Klepfisz, Daffern, & Day, 2016; Ward, 2016). Klepfisz et al. (2016) differentiate between actuarially useful risk factors and causal risk factors, highlighting that only those factors which reduce the likelihood of recidivism when changed can really be considered dynamic. Similarly, Mann, Hanson, and Thornton (2010) emphasize that there should be a focus in interventions on psychologically meaningful risk factors; that is, those psychological characteristics that are plausible possible causes of offending behaviour, and also reliably predict recidivism.

Identification of psychologically meaningful risk factors logically begins with examination of within treatment change and its expected impact on behaviour. Accordingly, psychometrically measured changes over the course of intervention have been considered informative in the assessment of reoffending risk. Therapeutic change has not consistently been found to be indicative of lower likelihood of future offending behaviour, but support has emerged for the use of measures of antisocial cognitive and social functioning (including hostility and impulsivity) to predict recidivism (Serin, Lloyd, Helmus, Derkzen, & Luong, 2013). However, the application of within treatment change evaluation to release and additional treatment or supervision planning currently depends on an insufficient empirical basis. Much research examining change data measures either magnitude of change, sometimes with comparison to a control group, or theoretical significance of change (comparability of pre- and post-treatment psychometric scores to non-offending normative data). This approach relies on an assumption that the constructs measured contain risk relevant information. In studies explicitly linking change to indices of recidivism, some research suggests that change in the desired direction (i.e., theoretically less problematic functioning) is indicative of reduced future risk. Overall, however, the results are mixed when examining the association of proximal assessments of treatment progress with distal outcomes.

For example, insight into anger problems and knowledge of and self-confidence in anger management skills have been associated with lower likelihood of reconviction for violent crime (Dowden, Blanchette, & Serin, 1999). In an evaluation of outcomes for 82 participants in a programme for violent offenders, significant change in the desired direction on a victim empathy measure, but not on a measure of minimization, predicted violent recidivism (O’Brien & Daffern, 2017). Kroner and Yessine (2013) found that of
nine scales measuring criminal attitudes and association with antisocial others, only change on one measure of antisocial association predicted recidivism. Measuring pre- to post-treatment change and behavioural outcomes among participants in the persistently violent offender programme, which ran in Canada until 2001, compared to participants in an anger and emotion management programme, and an attrition (from either programme) group, Serin, Gobeil, and Preston (2009) found no significant differences between groups on measures of treatment targets, institutional misconduct, nor post-release returns. Note that in this study, between-group pre- to post-treatment differences on the psychometric measures were non-significant, leaving untested the hypothesis that there is an association between change and recidivism. In another study, none of the change data from scales measuring criminal thinking styles and anger were predictive of violent reoffending (Klepfisz et al., 2014). However, this was based on a particularly small sample ($N = 42$) and a low recidivism base rate (26%). Olver, Kingston, Nicholaichuk, and Wong (2014) examined well-established measures of cognitive distortions, aggression/hostility, empathy, loneliness, social intimacy, and acceptance of responsibility in a sample of 392 sexual offenders. For most scales assessed, where associations were found with recidivism, they tended to be weak with the exception that reductions in self-reported aggressive characteristics shown in scores on the Aggression Questionnaire (AQ; Buss & Perry, 1992) were predictive of decreased general recidivism.

As the within-treatment change research evolves so do analytical approaches, with important implications for interpretation of results. Beyond demonstration of statistically significant differences between pre- and post-treatment scores on selected measures, Reliable Change Index (RCI) and clinically significant change (CSC) criteria provide additional information indicating whether or not a given score exceeds measurement error and converges with scores from a relevant normative sample (e.g., Klepfisz et al., 2014; Morgan, Kroner, Mills, Bauer, & Serna, 2014; Nunes, Babchishin, & Cortoni, 2011; Wakeling, Beech, & Freemantle, 2013). However, as researchers using RCI-CSC methods have noted, the validity of cut-off scores determining whether a given score may be considered ‘functional’ is dependent on the suitability of available normative data, usually meaning a range of scores and reliability and validity coefficients derived from a non-offending sample. This is problematic for at least two reasons. First, the clinical significance of a ‘functional’ score is tied to the premise that scores on selected measures should reflect fewer difficulties in the construct of interest among non-offender than offender samples, which is not consistently found. For instance, Hornsveld, Muris, and Kraaimaat (2011) found higher scores on the Novaco Anger Scale and Provocation Inventory (Novaco, 1994) in a sample of students than a sample of forensic psychiatric patients, and Williams et al. (1996) found lower scores on the AQ among adult male offenders compared to normative data from a student sample. Second, there is insufficient evidence that a shift into the ‘functional’ range necessarily results in behavioural change (Olver & Wong, 2013). Further, CSC status is likely to be confounded with risk, such that higher pre-treatment risk may explain poorer recidivism outcomes for offenders who improve in treatment but continue to be considered ‘dysfunctional’ based on post-treatment scores above the normative sample cut-off, compared to their peers with equal degree of change managing to move into the ‘functional’ range and have lower rates of recidivism by virtue of their lower pre-treatment risk (Olver, Beggs Christofferson, & Wong, 2015). The RCI-CSC method was therefore not employed for the purposes of this study in favour of an alternative approach described in the analytical strategy.
Rationale for the current study

Only a few studies have made explicit links between treatment change and recidivism while also statistically controlling for pre-treatment levels of risk and need (Klepfisz et al., 2014; Kroner & Yessine, 2013; O’Brien & Daffern, 2017; Olver et al., 2014). Some of this work focused on sexual offenders, and only a selection of commonly used psychometric instruments has been examined. Therefore, a larger evidence base determining risk relevance in psychometric tools is needed particularly in the non-sexual violence domain. To address the relative lack of research with violent offenders in this area of forensic clinical practice, the present study aimed to contribute to the psychometric risk relevance literature applicable to male offenders with a history of persistent violence. Given that maladaptive response to anger is an important dynamic factor, although but one of a number of others, associated with violent behaviour (Howells, 2004; Novaco, 2011), specific aims of the study were to examine the relationship between change in anger and aggressive response style measures in particular and recidivism. We also examined impulsivity, as a factor previously indicated to be relevant to the prediction of recidivism among violent offenders (Serin et al., 2013). Finally, we examined change in dynamic risk factors as evaluated using a structured quantitative approach. In the present study, we hypothesized associations of greater within-treatment change to be predictive of decreased recidivism, in line with previous findings that pre-treatment levels of risk and need would account for at least some observed individual differences in outcome.

Method

Sample

This study draws on data that were routinely collected in relation to the Violence Prevention Program (VPP), Correctional Service of Canada. The VPP is a cognitive behavioural programme aiming to reduce violence by enhancing motivation before targeting violence awareness, anger control, problem-solving, social attitudes, conflict resolution, self-control, and promoting positive lifestyle and relationships. Referral criteria are that participants should have been deemed persistently violent, meaning that they had committed a minimum of two violent offences and had been assessed to present high risk to commit future violent crimes, based on the Statistical Information on Recidivism Scale (SIR-R1; Nafekh & Motiuk, 2002), or the Offender Intake Assessment (Motiuk, 1997), which is an evaluation system used by the Correctional Services of Canada for planning purposes, that categorizes offenders as low, medium, or high risk through an assessment of static criminal risk factors. Outcome evaluation for the VPP suggests that, overall (noting that analyses have been conducted for both violent and nonviolent recidivism outcomes as well as for Indigenous and non-Indigenous subgroups), participants recidivate at a lower rate than non-participants (Higgs, Cortoni, & Nunes, 2019).

At the end of the study (July 31, 2011), follow-up data had been collated for 345 male federal offenders who had participated in the VPP between 1999 and 2004, including 188 non-completers. The present sample comprises those participants who had completed the programme and had therefore completed pre- and post-treatment assessments (N = 227). The mean sentence length that they had received was 6.7 years (SD = 6.64), and before the end of the study follow-up period, all of the participants had been released into the community. The average time that an offender was in the community following release, that is from their release date to either the recidivism date (date of revocation or
new offence) or the study follow-up period end date, was 3.31 years ($M = 1207.62$ days; range = 3 to 3736). The average age (at time of release) of the participants was 34.01 years ($SD = 8.75$).

**Measures**

**Aggression Questionnaire**
The Aggression Questionnaire (AQ; Buss & Perry, 1992) scores 29 items on a 5-point Likert-type scale. Higher scores are indicative of greater aggressiveness. The psychometric properties of the AQ have been examined quite extensively with results generally consistent with Buss and Perry’s (1992) validity coefficients. For example, meta-analytic data suggest good overall internal consistency: $\alpha = .89$, total; .81, anger; .79, hostility; .83, physical aggression; and .68, verbal aggression (Babchishin et al., 2011). However, the original four-factor structure found by Buss and Perry (1992) has generally not been supported in subsequent research with forensic samples (for a review, see Pettersen, Nunes, & Cortoni, 2018); thus, we examined total scores rather than subscales in our analyses.

**Novaco Anger Scale**
The Novaco Anger Scale (NAS; Novaco, 1994) is divided into two parts, measuring anger and reaction to provocation. Part A comprises 48 items relating to cognitive, arousal, and behavioural domains of anger. Responses are given on a 3-point scale, with higher scores suggesting that an angry response style is more characteristic of the respondent. Part B uses a 4-point scale, scored in the same direction, and contains 25 items to measure the extent to which respondents experience anger in response to perceived disrespectful treatment, unfairness/injustice, frustration/interruption, the annoying traits of others, and irritations. Good psychometric properties have been demonstrated including in offender samples. Mills, Kroner, and Forth (1998) reported internal consistency ($\alpha = .95$ and .97 for Part A and Part B, respectively), and test–retest reliability ($r = .78–.91$) in male general and violent offenders, and Baker, Van Hasselt, and Sellers (2008) found reliability in male domestic violence or drug-related offenders ($\alpha = .93$ and .92 for Part A and Part B, respectively) as well as convergent validity based on correlation with other measures of anger including total score on Spielberger, Jacobs, Russell, and Crane’s (1983) State-Trait Anger Scale ($r = .61$ and .50 for Part A and Part B, respectively).

**I7-Impulsivity Scale**
The Impulsivity Scale (I7-Impulsivity) comprises 19 items of the 54-item Impulsiveness-venturesomeness-empathy questionnaire (Eysenck, Pearson, Easting, & Allsopp, 1985), dichotomously scored (yes/no), with higher scores suggesting greater impulsivity. Babchishin et al. (2011) found good internal consistency ($\alpha = .83$), and Gordon and Egan (2011) recommend the I7-Impulsivity Scale over other commonly used self-report measures of impulsivity in forensic settings, based on associations found between scale scores and retrospective criminal conviction and prison breach data (AUC = .67, 95% CI 0.58–0.81 for violent criminal convictions).

**Violence Risk Scale**
The Violence Risk Scale (VRS; Wong & Gordon, 1999–2003) is a violence risk assessment and treatment planning tool designed to assess risk for violence, identify targets for
treatment, and assess change in risk from violence reduction treatment or other change agents. The VRS is comprised of six static (i.e., historic, generally unchanging) and 20 dynamic (i.e., potentially changeable) items, each of which are rated on a four-point (0, 1, 2, and 3) ordinal scale. Higher item scores indicate a positive association with increased risk for violence. The items are summed to yield static, dynamic, or total (static + dynamic) scores. Changes on the VRS are assessed by way of a modified application of the stages of change (SoC) model of the change process arranged into five stages: pre-contemplation, contemplation, preparation, action, and maintenance. The five stages are operationalized for each of the 20 dynamic items documenting the cognitive, experiential, and behavioural changes that occur as an individual remediates problem areas towards reducing their violence risk. Items with a 2 or 3 rating (denoting criminogenic needs) are given a baseline SoC at pre-treatment, and then, the SoC is rerated at post-treatment. Progression from one stage to the next, denoting progress, corresponds to a 0.5 deduction in score, two stages 1.0 point, and so on; regressing a stage in the case of deterioration is given a corresponding 0.5 increase. Item change ratings are summed to generate a change score, which is deducted from (or, in the case of deterioration, added to) the pre-treatment dynamic score to yield a post-treatment score. Psychometric research on the VRS supports the predictive accuracy of risk (O’Brien & Daffern, 2017; Wong & Gordon, 2006) and change scores (Lewis, Olver, & Wong, 2013) for future violence. Therefore, in the present study, VRS scores were used to (1) evaluate the risk relevance of the psychometric measures through their convergent validity with a measure of risk for future violence and (2) test the incremental associations of change when controlling for baseline risk. The VRS was rated in real time by trained VPP personnel pre and post-treatment.

Post-release outcomes
Outcome measurement focused on recidivism data obtained from Canadian Police Information Centre records, Canada’s national criminal record database maintained by the Royal Canadian Mounted Police, which contain the full history of criminal charges, convictions, and dispositions for all offenders in Canada. Recidivism variables were operationalized as any new criminal conviction for a new offense incurred post-release (in contrast to any new charge, which may not have reached conviction). Recidivism variables were coded dichotomously (occurrence or not) according to criminal code for any violent (sexual or nonsexual) offense, or general (i.e., any new conviction be it violent or nonviolent) offense, along with date of reconviction.

Analytic strategy
An initial screening showed that data were incomplete for some measures. Little’s MCAR test suggested that data were missing completely at random; therefore, for each scale, participants with less than 10% of their responses missing were identified and missing values were replaced with the series mean (except for the VRS, for which prorating rules in accordance with the VRS scoring manual were applied), in order to minimize the impact of missing data on statistical power. This resulted in imputing data for 3–8 (1.3–3.5%) of the 227 participants per scale. Remaining missing data were managed in analysis through pairwise exclusion. Pre- to post-treatment change scores (pre-scores minus post-scores) were calculated for all participants with valid scores on each of the psychometric and risk measures.
Scores on the psychometric scales (AQ; NAS Part A; NAS Part B; I7-Impulsivity) and pre- to post-treatment changes were then subjected to a series of tests of risk relevance. First, Pearson’s correlation was used to test the association between the scales and estimated recidivism risk, measured using the VRS. Next, receiver operating characteristic (ROC) analysis was used to assess the predictive accuracy of change for each measure. The magnitude of the area under the curve (AUC) generated by ROC analysis corresponds to small, medium, and large effects for values of .56, .63, and .71, respectively (Rice & Harris, 2005). AUCs were obtained using residualized change scores, which offer a statistical correction for variance in the change score accounted for by the pre-treatment score (Beggs & Grace, 2011), and are therefore preferable to RCI-CSC methods. Residualized change was calculated by regressing the raw change score for each measure on the pre-treatment score, so that residualized change is equal to actual change minus predicted change based on pre-treatment score. Finally, change predictors were selected for logistic regressions on fixed 3- and 5-year recidivism outcomes based on significant univariate associations with recidivism, controlling for baseline risk (i.e., VRS pre-treatment score).

Results

Descriptive statistics and change associations with violence risk
In keeping with the VPP’s mandate, this was a high risk-need sample with a mean VRS score corresponding to the high risk range (i.e., 50+), and approximately one full standard deviation above the mean for the normative sample (Wong & Gordon, 2006). On average, the men changed by nearly a full standard deviation on each of the measures from pre- to post-treatment (see Table 1). As shown in Table 1, scores on the AQ, NAS Part A, NAS Part B, and the I7-Impulsivity scales correlated positively with VRS scores, but the associations were generally weak.

Recidivism
The sample was followed up for 7.0 years ($SD = 2.2$) post-release, during which, 40.1% (91/227) were convicted for a new violent offense and 65.2% (148/227) for any new offense. Recidivism outcomes were examined using fixed 3- and 5-year follow-ups to control for follow-up time; rates of violent recidivism were 23.1% (50/216) and 37.2% (73/196) for these follow-ups, respectively, while rates of general recidivism were 43.5% (94/216) and 58.2% (114/196), respectively.

Changes on core psychological domains and associations with recidivism
To test the individual risk relevance of each of the scales, the degree to which pre- and post-treatment scores predicted likelihood of recidivism was assessed using a series of ROC analyses. In order to evaluate magnitude of change on each of the scales, while controlling for pre-treatment score, predictive accuracy of residualized change scores was tested (Table 2). AUC values indicated that pre- to post-treatment changes on the psychometric test scores were a poor predictor for both types of recidivism outcome (reconviction for any new offenses, and reconviction for violent offenses). Only the VRS dynamic and AQ change scores predicted decreased recidivism at $p < .05$. This was only true for general recidivism, not violent recidivism. Predictive accuracy generally decreased over time, with weaker AUCs for most of the measures at 5-year follow-up.
than at 3-year follow-up (as may be reasonably anticipated, since both recidivism status and indicators, that is psychometric scores, are values that may change over time, decreasing AUCs are typical the further the prediction time is from the time of indicator measurement; Kamarudin, Cox, & Kolamunnage-Dona, 2017).

Subsequently, the incremental associations of VRS and AQ measured change were tested using logistic regressions, controlling for baseline risk by entering the VRS pre-treatment total (static and dynamic) score in the first block of each model (four models were tested for VRS residualized change and for AQ residualized change, and each recidivism outcome, general and violent, at the two fixed follow-up times). Classification rates and model fit were compromised by overclassification into the largest group: recidivists in the general recidivism at 5-year model and non-recidivists in the other three models. For general recidivism, the overall correct classification rate was 61% and 57% at 3 and 5 years, respectively (Nagelkerke $R^2 = .06$ and .05). For violent recidivism, the correct classification rate was 78% and 64% at 3 and 5 years, respectively (Nagelkerke $R^2 = .01$ and .001). Table 3 shows the contribution of the individual predictors, regression coefficients, odds ratios, and the 95% confidence intervals around them. For general recidivism, the probability of recidivism was lower for VPP participants showing greater change on the VRS, but not the AQ, at both 3 and 5 years. For violent recidivism, incremental associations of measured change were non-significant at 3 years as well as at 5 years.

Table 1. Descriptive statistics and correlations between the Aggression Questionnaire (AQ), Impulsivity Scale (I7), and Novaco Anger Scale (NAS Part A; NAS Part B), and the Violence Risk Scale (VRS) pre-treatment, post-treatment, and change scores for Violence Prevention Program participants

| Measure              | VRS (r) | Pre-treatment | Post-treatment | Change | $M$ (SD) | d change |
|----------------------|---------|---------------|----------------|--------|----------|----------|
| AQ pre               | .13     | .14           | -.01           |        | 83.2 (21.5) | —        |
| AQ post              | .07     | .15           | -.15           |        | 63.3 (17.5) | —        |
| AQ change            | .10     | .13           | .12            |        | 19.9 (21.9) | 1.01**   |
| I7 pre               | .11     | .09           | .06            |        | 8.8 (4.6) | —        |
| I7 post              | .03     | .14           | -.20*          |        | 5.2 (4.0) | —        |
| I7 change            | .04     | -.03          | .20*           |        | 3.6 (4.8) | .83**    |
| NAS Part A pre       | .12     | .05           | .12            |        | 85.0 (15.5) | —        |
| NAS Part A post      | .08     | .07           | -.02           |        | 71.5 (13.0) | —        |
| NAS Part A change    | .10     | .11           | .13            |        | 13.5 (17.7) | .94**    |
| NAS Part B pre       | .13     | .12           | .05            |        | 53.5 (14.0) | —        |
| NAS Part B post      | .07     | .05           | .01            |        | 43.0 (11.5) | —        |
| NAS Part B change    | .11     | .17           | .02            |        | 10.5 (15.2) | .81**    |
| VRS static           | —       | —             | —              |        | 13.2 (2.8) | —        |
| VRS dynamic pre      | —       | —             | —              |        | 44.3 (8.0) | —        |
| VRS dynamic post     | —       | —             | —              |        | 36.7 (7.1) | —        |
| VRS change           | —       | —             | —              |        | 7.5 (4.0) | .97**    |
| VRS total pre        | —       | —             | —              |        | 57.5 (9.0) | —        |
| VRS total post       | —       | —             | —              |        | 50.0 (8.2) | —        |

Notes. AQ $n = 167$, I7 $n = 161$, NAS Part A $n = 156$, NAS Part B $n = 152$, VRS $n = 134$.

*p $\leq .05$, **p $\leq .001$. 

Table 1. Descriptive statistics and correlations between the Aggression Questionnaire (AQ), Impulsivity Scale (I7), and Novaco Anger Scale (NAS Part A; NAS Part B), and the Violence Risk Scale (VRS) pre-treatment, post-treatment, and change scores for Violence Prevention Program participants

| Measure              | Pre-treatment | Post-treatment | Change | $M$ (SD) | d change |
|----------------------|---------------|----------------|--------|----------|----------|
| AQ pre               | .13           | .14           | -.01   | 83.2 (21.5) | —        |
| AQ post              | .07           | .15           | -.15   | 63.3 (17.5) | —        |
| AQ change            | .10           | .13           | .12    | 19.9 (21.9) | 1.01**   |
| I7 pre               | .11           | .09           | .06    | 8.8 (4.6) | —        |
| I7 post              | .03           | .14           | -.20*  | 5.2 (4.0) | —        |
| I7 change            | .04           | -.03          | .20*   | 3.6 (4.8) | .83**    |
| NAS Part A pre       | .12           | .05           | .12    | 85.0 (15.5) | —        |
| NAS Part A post      | .08           | .07           | -.02   | 71.5 (13.0) | —        |
| NAS Part A change    | .10           | .11           | .13    | 13.5 (17.7) | .94**    |
| NAS Part B pre       | .13           | .12           | .05    | 53.5 (14.0) | —        |
| NAS Part B post      | .07           | .05           | .01    | 43.0 (11.5) | —        |
| NAS Part B change    | .11           | .17           | .02    | 10.5 (15.2) | .81**    |
| VRS static           | —             | —             | —      | 13.2 (2.8) | —        |
| VRS dynamic pre      | —             | —             | —      | 44.3 (8.0) | —        |
| VRS dynamic post     | —             | —             | —      | 36.7 (7.1) | —        |
| VRS change           | —             | —             | —      | 7.5 (4.0) | .97**    |
| VRS total pre        | —             | —             | —      | 57.5 (9.0) | —        |
| VRS total post       | —             | —             | —      | 50.0 (8.2) | —        |
The simple endorsement of significantly fewer problem indicative items on a measure of interest by an individual after completing an intervention has been demonstrated to be insufficient to inform treatment progress and subsequent assessments of risk of future offending (e.g., Klepfisz et al., 2014; Olver et al., 2014; Wakeling et al., 2013). The present study aimed to contribute to the limited extant literature providing more stringent tests of the meaningfulness of pre- to post-treatment changes on measures widely used in the assessment of violent offenders. Results were consistent with recent developments in research examining psychometric data and its risk relevance; it may be erroneous to rely on apparent within-treatment change in some key constructs – such as anger and impulsivity – ostensibly targeted and measured over the course of intervention when tasked with offenders’ ongoing sentence and release planning. That is, self-report measures of these constructs do not permit inferences about risk-related behavioural change. Conversely, changes in VRS dynamic score being associated with decreased general (although not violent) recidivism in a conservative test of risk relevance.

### Table 2. Predictive accuracy of pre-treatment, post-treatment, and residualized change (r. change) scores on psychometric and risk measures for general and violent recidivism at 3- and 5-year fixed follow-up

| Measure       | General recidivism | Violent recidivism |
|---------------|--------------------|--------------------|
|               | 3 years            | 5 years            | 3 years            | 5 years            |
|               | AUC 95% CI         | AUC 95% CI         | AUC 95% CI         | AUC 95% CI         |
| AQ pre        | .55 (0.46, 0.64)   | .59 (0.50, 0.69)   | .44 (0.34, 0.55)   | .51 (0.41, 0.61)   |
| AQ post       | .64** (0.55, 0.72) | .59 (0.50, 0.68)   | .54 (0.45, 0.64)   | .49 (0.39, 0.58)   |
| AQ r. change  | .62** (0.53, 0.71) | .56 (0.47, 0.65)   | .56 (0.46, 0.66)   | .48 (0.38, 0.58)   |
| I7 pre        | .52 (0.43, 0.61)   | .54 (0.45, 0.64)   | .51 (0.41, 0.61)   | .49 (0.39, 0.59)   |
| I7 post       | .57 (0.48, 0.66)   | .55 (0.46, 0.65)   | .54 (0.44, 0.64)   | .48 (0.39, 0.58)   |
| I7 r. change  | .56 (0.46, 0.65)   | .53 (0.43, 0.62)   | .52 (0.41, 0.63)   | .47 (0.37, 0.57)   |
| NAS Part A pre | .47 (0.37, 0.56) | .41 (0.31, 0.51) | .43 (0.31, 0.54) | .42 (0.32, 0.52) |
| NAS Part A post | .44 (0.35, 0.54) | .42 (0.32, 0.52) | .44 (0.33, 0.55) | .44 (0.34, 0.54) |
| NAS Part A r. change | .45 (0.35, 0.54) | .44 (0.34, 0.54) | .46 (0.34, 0.57) | .47 (0.37, 0.57) |
| NAS Part B pre | .47 (0.38, 0.57) | .39* (0.29, 0.48) | .51 (0.40, 0.63) | .48 (0.37, 0.59) |
| NAS Part B post | .50 (0.40, 0.60) | .43 (0.33, 0.53) | .50 (0.39, 0.61) | .48 (0.38, 0.59) |
| NAS Part B r. change | .50 (0.41, 0.59) | .45 (0.35, 0.55) | .49 (0.37, 0.61) | .49 (0.39, 0.59) |
| VRS static    | .56 (0.46, 0.66)   | .59 (0.49, 0.70)   | .56 (0.44, 0.68)   | .53 (0.42, 0.64)   |
| VRS dynamic pre | .47 (0.37, 0.57) | .45 (0.35, 0.56) | .43 (0.31, 0.55) | .47 (0.36, 0.58)   |
| VRS dynamic post | .56 (0.46, 0.66) | .54 (0.43, 0.64) | .46 (0.33, 0.59) | .46 (0.35, 0.58)   |
| VRS r. change | .61* (0.51, 0.71)  | .59 (0.49, 0.70)   | .54 (0.41, 0.66)   | .50 (0.39, 0.61)   |
| VRS total pre | .49 (0.38, 0.59)   | .49 (0.39, 0.60)   | .45 (0.33, 0.57)   | .47 (0.36, 0.58)   |
| VRS total post | .57 (0.47, 0.67)  | .56 (0.46, 0.67)   | .48 (0.36, 0.61)   | .48 (0.37, 0.59)   |

**Notes.** For change score analyses, recidivism outcome is reverse-coded such that positive AUC values represent associations between positive change and decreased recidivism. AQ = Aggression Questionnaire, n = 161 (3 years), n = 147 (5 years); I7 = Impulsivity Scale, n = 155 (3 years), n = 141 (5 years); NAS = Novaco Anger Scale, n = 145 (3 years), n = 133 (5 years); VRS = Violence Risk Scale, n = 128 (3 years), n = 119 (5 years).

*p ≤ .05, **p ≤ .01.

### Discussion

The simple endorsement of significantly fewer problem indicative items on a measure of interest by an individual after completing an intervention has been demonstrated to be insufficient to inform treatment progress and subsequent assessments of risk of future offending (e.g., Klepfisz et al., 2014; Olver et al., 2014; Wakeling et al., 2013). The present study aimed to contribute to the limited extant literature providing more stringent tests of the meaningfulness of pre- to post-treatment changes on measures widely used in the assessment of violent offenders. Results were consistent with recent developments in research examining psychometric data and its risk relevance; it may be erroneous to rely on apparent within-treatment change in some key constructs – such as anger and impulsivity – ostensibly targeted and measured over the course of intervention when tasked with offenders’ ongoing sentence and release planning. That is, self-report measures of these constructs do not permit inferences about risk-related behavioural change. Conversely, changes in VRS dynamic score being associated with decreased general (although not violent) recidivism in a conservative test of risk relevance.
Table 3. Logistic regression: incremental associations of VRS and AQ measured change (residualized scores) with fixed 3- and 5-year general and violent recidivism controlling for baseline risk (VRS pre-treatment score)

| Regression model (1-4) | 3-year recidivism | 5-year recidivism |
|------------------------|-------------------|-------------------|
|                        | B     | SE   | Wald | p     | e^B   | 95% CI    | B     | SE   | Wald | p     | e^B   | 95% CI    |
| General recidivism     |       |      |      |       |       |           |       |      |      |       |       |           |
| 1                      | VRS pre total    | .031  | .022 | 1.98  | .160  | 1.032     | .024  | .024 | 1.03 | .311  | 1.024 | 0.978, 1.072 |
|                        | VRS res. change  | -.116 | .052 | 5.03  | .025  | 0.890     | -.104 | .052 | 3.90 | .048  | 0.902 | 0.814, 0.999 |
| 2                      | VRS pre total    | .012  | .024 | 0.24  | .623  | 1.012     | .018  | .024 | 0.54 | .462  | 1.018 | 0.971, 1.068 |
|                        | AQ res. change   | -.022 | .012 | 3.20  | .074  | 0.979     | -.002 | .013 | 0.02 | .891  | 0.998 | 0.973, 1.024 |
| Violent recidivism     |       |      |      |       |       |           |       |      |      |       |       |           |
| 3                      | VRS pre total    | -.003 | .025 | 0.01  | .920  | 0.997     | -.005 | .023 | 0.05 | .818  | 0.995 | 0.950, 1.041 |
|                        | VRS res. change  | -.040 | .059 | 0.45  | .501  | 0.961     | .000  | .052 | 0.00 | .995  | 1.000 | 0.904, 1.107 |
| 4                      | VRS pre total    | .000  | .028 | 0.00  | .992  | 1.000     | .000  | .026 | 0.00 | .989  | 1.000 | 0.950, 1.052 |
|                        | AQ res. change   | -.014 | .014 | 1.01  | .316  | 0.986     | .012  | .014 | 0.80 | .370  | 1.012 | 0.985, 1.040 |

Notes. VRS change analyses, n = 128 (3 years), n = 99 (5 years); AQ change analyses, n = 106 (3 years), n = 99 (5 years). Significant p-values in bold font.
favours the structured actuarial approach over psychometric assessment in the evaluation of dynamic risk factors.

**Associations between within-treatment change and recidivism**

In the present study, pre- to post-treatment change on the self-report scales assessing components of anger and impulsivity (the Aggression Questionnaire, the I7-Impulsivity Scale, and the Novaco Anger Scale, Part A and Part B) generally did not predict recidivism for persistently violent offenders. The only exception was that decreases on the AQ significantly predicted lower general recidivism with the 3-year follow-up. Therefore, consistent with previous research (Olver *et al.*, 2014), results for the AQ were more promising than for other measures; however, in the present study, the AQ did not predict recidivism when baseline risk as well as variance in the change score accounted for by the pre-treatment score were controlled.

Moreover, only the VRS maintained significant associations with reductions in recidivism after controlling for baseline risk, but changes on the VRS were significantly associated with decreased general, but not violent, recidivism. Given that the VRS is a purpose-built dynamic risk measure, and the men attended a violence reduction programme, it makes sense that changes on the VRS would be risk relevant. It is unclear why similar associations for the VRS were not observed with respect to general violence in the present study, although one possibility may be that programme-related changes tend to reduce risk in a broader, general sense (e.g., improved problem solving, decreased antisocial attitudes) but that have less of a specific impact on violence. Indeed, prior research has found changes on the VRS linked to violence reduction programming to also be associated with decreased general recidivism (e.g., Coupland & Olver, 2018; Lewis *et al.*, 2013).

By contrast, with few exceptions, the self-report measures included in the present study were unreliable markers of risk-related progress. Either the measures failed to capture the constructs of interest relevant to offending behaviour, or the presumed constructs themselves were not psychologically meaningful risk factors as defined by Mann *et al.* (2010). It has been argued that the concept of dynamic risk factors, and even the concept of psychologically meaningful risk factors which assumes a causal relationship with offending behaviour, is fundamentally flawed when applied in explanatory efforts as well as in the prediction of offending. In this line of thinking, dynamic risk factors are predictive constructs, not explanatory concepts (Ward, 2016). In critiquing the assumption that dynamic risk factors are or should be informative in terms of causal processes, Ward (2016) asserts that a relationship has not been established between indicators, such as psychometric scores, of particular risk factors and the latent construct, dynamic risk factors. Although further research is clearly needed, the outcome of the present study echoes this: change on dynamic risk factors, measured by the VRS, had risk relevance but psychometric scores were not good indicators of risk-related change.

Alternatively, the association between anger and violent offending (Serin *et al.*, 2013) may be indirect or may apply to certain violent offenders differently than to others. Anger (likewise, impulsivity) is often elevated among offenders and commonly occurs in the presence of high psychopathic traits (Decuyper, De Pauw, De Fruyt, De Bolle, & De Clercq, 2009; Suter *et al.*, 2002). Yet, anger does not necessarily characterize violent offenders differently from non-violent offenders and thus may represent an erroneous (non-criminogenic) treatment target despite an offender’s history of violence and
aggression (Loza & Loza-Fanous, 1999). For example, compared to other offenders, those with high levels of psychopathic traits are more likely to violently reoffend (Campbell, French, & Gendreau, 2009). Yet, depending on whether psychopathic traits predominantly relate to interpersonal/affective characteristics as opposed to lifestyle/antisociality, violent acts may be callous and unemotional rather than expressive in nature (Tew, Harkins, & Dixon, 2013). Similarly, even when anger is identified as a relevant predisposing factor in a given case, the precise role of anger as a precursor to violence may not be clear (Stefanska et al., 2018). In the present sample, participants were persistently violent offenders. However, the data necessary to explore hypotheses relating to the potential influence of psychopathic traits were unavailable and the proportion of instrumental versus expressive violent acts could not be determined. Among offenders using instrumental violence, improvements seen in measures of anger and impulsivity may be unrelated to likelihood of reoffending, whereas it is possible that risk relevance would be found for the measures examined in a sample known to have typically used violence reactively. As such, the present research does not necessarily undermine offending behaviour programmes with anger control as a core component; indeed, changes in violence risk measured by the VRS predicted decreases in general recidivism. Instead, findings suggest that further research is needed to establish whether the psychometric measures in question, and the constructs they purport to measure, predict recidivism for specific subgroups of violent offenders.

**Implications for change conceptualization and risk management**

Further interpretations of the present findings might include a hypothesis that genuine within-treatment changes pertaining to decreased violence risk were simply not maintained post-release, although this had relevance for general recidivism as discussed previously. Polaschek and Yesberg (2017) followed up violent offenders over a period of 12 months on parole, finding that their sample of treatment completers had higher protective and lower stable and acute dynamic risk factors (using Dynamic Risk Assessment for Offender Re-Entry; Serin, 2007) than their comparison sample when entering the community. Growth models suggested that the better risk status of the treatment completers was sustained in the community, although analysis of rate of change showed that both the treatment completers and the comparison sample continued to improve at a similar rate. In the present study, data were only collected directly prior to, and after the completion of the VPP. Therefore, it is unfortunately not possible to comment further on whether change was or was not sustained, and whether a maintenance problem may have impacted our results.

Additionally, offenders appearing more troubling to parole boards based on reports of institutional management issues (among other sources of information), may be those same offenders whose pre-treatment psychometric data suggested highly problematic functioning in domains related to treatment, and regardless of change demonstrated on psychometric measures, remain highly concerning to those tasked with release planning. For example, impulsive behaviour may be relatively easily observed and recorded by institutional staff, and although psychometrically assessed pre-/post-treatment improvement may be looked upon favourably, there may be a tendency towards more cautious release plans because of the greater attention that these offenders attract institutionally. Consequently, it is possible that recidivism outcomes may be associated with more or less restrictive supervision and monitoring plans for offenders whose institutional behaviour is more or less challenging or obviously problematic (as in the case of reactive aggression...
vs. overcontrolled anger) despite comparable overall future violence risk. This scenario would suggest that recidivism risk has less to do with internal positive change, or the lack of it, and more to do with external factors. However, planning and implementation of community supervision was not evaluated in the present study and, as such, these are merely hypotheses offered in critique of the present findings and are not meant as a commentary on issues in the parole decision process or the quality of community supervision.

**Study strengths, limitations, and conclusions**

Within-treatment change studies hold important implications for correctional service policy and practice, but there is no academic consensus as to the most effective methodological approach. The RCI and analysis of CSC have been recommended and can be seen to be practically and clinically advantageous given its intuitive individual-level utility (Klepfisz et al., 2014). On the other hand, the value of the CSC method has been disputed because of limitations associated with its dependence on the quality of the measures used and normative data, as well as a lack of independence from pre-treatment risk status (Olver et al., 2015; Wakeling et al., 2013). Here, residualized change scores were favoured as this method provides a statistical control for variance associated with pre-treatment scores, while also assessing change independently within the sample. That is, participants provided their own baseline rather than being dependent on cut-off scores determined by norms that may arbitrarily distinguish ‘functional’ and ‘dysfunctional’ groups. In terms of the study’s strengths and limitations and the treatment of the available data, it should be noted that statistical power was certainly negatively impacted by incomplete data and the reader’s attention is drawn not only to the overall relatively modest sample size, but also the number of participants included in the various statistical tests.

Time-series data extending into the community follow-up period would have allowed analysis of maintenance of change, which would have been beneficial and should be considered in future research with similar aims (see Polaschek & Yesberg, 2017, for what we believe to be the first study of this kind in the field of non-sexual violence research). Additionally, for the purposes of examining the risk relevance of psychometric measures, the present study focused on four scales measuring just two constructs, anger and impulsivity, whereas this fails to represent all components of multifactorial programmes expected to successfully reduce violent behaviour (Polaschek, 2006). It should also be noted that the present study tested the risk relevance of some of the psychometric measures commonly used in correctional services not to suggest that they might outperform multifactorial risk assessments, but to demonstrate their utility (or lack thereof) in informing treatment progress evaluations and as evidence to support decisions in rating relevant items within structured professional judgement risk assessment tools. Therefore, the implications of the study relate to the practice of using psychometric measures to infer behavioural change in dynamic risk factors rather than to methods of risk assessment per se. The results – the unfavourable findings for the psychometric scales and the superior relative performance of the VRS dynamic scores – suggest that the use of an established risk-need measure of dynamic risk may be more relevant to assess risk-related change than classic psychometric instruments.

Even though there is a certain degree of conceptual overlap between the psychometric scales assessed and some VRS dynamic items, the present findings support the monitoring of change using the VRS rather than through psychometric evaluations. This may be due to
factors such as the VRS scoring procedures which require not only an evaluation of the presence or absence of risk factors, but also the individual’s stage of change according to a modification of the transtheoretical model (Prochaska & DiClemente, 1984; Prochaska, DiClemente, & Norcross, 1992). Indeed, progression across the stages reflects the use of skills and strategies to manage a given problem area captured by the item. In addition, it is possible that the VRS dynamic score simply offered a more comprehensive and complete picture of change across the 20 dynamic risk factors compared to anger and impulsivity as specific indicators. In this study, risk relevance may also have been differentially associated with Factor 1 of the VRS (Wong & Gordon, 2006) which includes ten dynamic items including interpersonal aggression and impulsivity as well as violent lifestyle, emotional regulation/control, violence during institutionalization, weapon use, substance abuse, stability of relationships, violence cycle, and cognitive distortion; compared to Factor 2 (one dynamic item: criminal peers); or Factor 3 (nine dynamic items: criminal personality, criminal attitude, work ethic, insight into violence, community support, release to high-risk situations, compliance with supervision, security level of release institution, and cognitive distortions, which cross-loads; Wong & Gordon, 2006). These remain questions for further research.

Further, we have avoided suggestion that the present findings speak to the effectiveness of the VPP. It should be made explicit that in the absence of a control group, the present study is limited in its capacity to answer the question as to what extent change on psychometric assessments may be attributed to programme participation. Indeed, the study was concerned with the risk relevance of change, but not the source of that change (i.e., whether attributable to treatment or other agents). For the VRS, however, it may be reasonable to attribute the changes, at least in part, to treatment since it was rated at two time points, with change explicitly assessed in relation to progress in programming.

Overall, it is clear that further research is needed to elucidate the mechanisms of change that may be activated through interventions with violent offenders to facilitate their successful reintegration into the community. Important inroads have been achieved through formal examination of dynamic risk instruments, such as the VRS (Coupland & Olver, 2018; Lewis et al., 2013; O’Brien & Daffern, 2017), reassessed at multiple timepoints, which are intended to measure and track progress towards risk reduction. Elsewhere, mathematical models have been employed to integrate risk and change information so that actuarial estimates can be adjusted to reflect changes in risk in a systematic and non-arbitrary manner with sexual offenders (Olver et al., 2018). The next step, as we see it, is to extend such applications to violent offending populations.

Conflicts of interest
All authors declare no conflict of interest.

Author contribution
Tamsin Higgs, DForenPsy. (Conceptualization; Formal analysis; Methodology; Writing – original draft; Writing – review & editing) Mark E Olver (Formal analysis; Methodology; Writing – review & editing) Kevin L Nunes (Writing – review & editing) Franca Cortoni (Conceptualization; Data curation; Supervision)
Data Availability Statement

Research data are not shared.

References

Andrews, D. A., & Bonta, J. L. (2010a). The psychology of criminal conduct. Cincinnati, OH: Anderson.
Andrews, D. A., & Bonta, J. L. (2010b). Rehabilitating criminal justice policy and practice. Psychology, Public Policy, and Law, 16(1), 39–55. https://doi.org/10.1037/a0018362
Babchishin, K. M., Pettersen, C., Nunes, K. L., & Cortoni, F. (2011). Meta-analysis of functional and dysfunctional norms for measures of treatment targets for violence offender programs. In F. Cortoni (Chair), Reducing recidivism among violent offenders: The effectiveness of treatment. Paper presented at the Second North American Correctional and Criminal Justice Psychology Conference, Toronto, Ontario.
Baker, M., Van Hasselt, V. B., & Sellers, A. H. (2008). Validation of the Novaco Anger scale in an incarcerated offender population. Criminal Justice and Behavior, 35, 741–754. https://doi.org/10.1177/0093854808316275
Beggs, S. M., & Grace, R. C. (2011). Treatment gains for sexual offenders against children predicts recidivism: A comparative validity study. Journal of Consulting and Clinical Psychology, 79, 182–192. https://doi.org/10.1037/a0022900
Buss, A. H., & Perry, M. (1992). The Aggression Questionnaire. Journal of Personality and Social Psychology, 63, 452–459. https://doi.org/10.1037/002-3514.63.3.452
Campbell, M. A., French, S., & Gendreau, P. (2009). The prediction of violence in adult offenders: A meta-analytic comparison of instruments and methods of assessment. Criminal Justice and Behavior, 36, 567–590. https://doi.org/10.1177/009385480933610
Coupland, R. B. A., & Olver, M. E. (2018). Assessing dynamic violence risk in a high risk treated sample of violent offenders. Assessment, 1–15. Advance online publication. https://doi.org/10.1177/1073191118797440
Decuyper, M., De Pauw, S., De Fruyt, F., De Bolle, M., & De Clercq, B. J. (2009). A meta-analysis of psychopathy-, antisocial PD- and FFM associations. European Journal of Personality, 23, 531–565. https://doi.org/10.1002/per.729
Dowden, C., Blanchette, K., & Serin, R. (1999). Anger management programming for federal male inmates: An effective intervention (R-82). Retrieved from www.csc-scc.gc.ca/research
Eysenck, S. B. G., Pearson, P. R., Easting, G., & Allsopp, J. F. (1985). Age norms for impulsiveness, venturesomeness and empathy in adults. Personality and Individual Differences, 6, 613–619.
Gordon, V., & Egan, V. (2011). What self-report impulsivity measure best postdicts criminal convictions and prison breaches of discipline? Psychology, Crime & Law, 17, 305–318. https://doi.org/10.1080/10683160903203946
Hanson, R. K., Bourgon, B., Helmus, L., & Hodgson, S. (2009). The principles of effective correctional treatment also apply to sexual offenders: A meta analysis. Criminal Justice and Behavior, 36, 865–891. https://doi.org/10.1177/0093854809338545
Henwood, K. S., Chou, S., & Browne, K. D. (2015). A systematic review and meta-analysis on the effectiveness of CBT informed anger management. Aggression and Violent Behavior, 25, 280–292. https://doi.org/10.1016/j.avb.2015.09.011
Higgs, T., Cortoni, F., & Nunes, K. (2019). Some positive recidivism outcomes for Canadian treated high-risk offenders. Criminal Justice and Behavior, 46, 359–373. https://doi.org/10.1177/0093854818808850
Hornsved, R. H., Muris, P., & Kraaimaat, F. W. (2011). The Novaco Anger Scale-Provocation Inventory (1994 version) in Dutch forensic psychiatric patients. Psychological Assessment, 23, 937–944. https://doi.org/10.1037/a0024018
Howells, K. (2004). Anger and its links to violent offending. Psychiatry, Psychology and Law, 11, 189–196. https://doi.org/10.1375/1321871042707278
Implications for risk-change communication. *Behavioral Sciences and the Law, 33*, 92–110. https://doi.org/10.1002/bls.2159

Olver, M. E., Kingston, D. A., Nicholaichuk, T. P., & Wong, S. C. P. (2014). A psychometric evaluation of treatment change in a multisite sample of treated Canadian federal sexual offenders. *Law and Human Behavior, 38*, 544–599. https://doi.org/10.1037/lhb0000086

Olver, M. E., Mundt, J. C., Thornton, D., Beggs Christofferson, S. M., Kingston, D. A., Sowden, J. N., Nicholaichuk, T. P., Gordon, A., & Wong, S. C. P. (2018). Using the Violence Risk Scale-Sexual Offense Version in sexual violence risk assessments: Updated risk categories and recidivism estimates from a multisite sample of treated sexual offenders. *Psychological Assessment, 30*, 941–955. https://doi.org/10.1037/pas0000538

Olver, M. E., & Wong, S. C. P. (2013). Treatment programs for high risk sexual offenders: Program and offender characteristics, attrition, treatment change and recidivism. *Aggression and Violent Behavior, 18*, 579–591. https://doi.org/10.1016/j.avb.2013.06.002

Palmer, T. (1975). Martinson revisited. *Journal of Research in Crime and Delinquency, 12*, 133–152. https://doi.org/10.1177/002242787501200206

Pettersen, C., Nunes, K. L., & Cortoni, F. (2018). The factor structure of the Aggression Questionnaire with violent offenders. *International Journal of Offender Therapy and Comparative Criminology, 62*, 1888–1905. https://doi.org/10.1177/0306624X17691242

Polaschek, D. L. L. (2006). Violent offender programmes: Concept, theory, and practice. In C. R. Hollin & E. J. Palmer (Eds.), *Offending behaviour programmes: Development, application and controversies* (pp. 113–154). Chichester, UK: John Wiley & Sons.

Polaschek, D. L. L., & Yesberg, J. A. (2017). High-risk violent prisoners’ patterns of change on parole on the DRAOR’s dynamic risk and protective factors. *Criminal Justice and Behavior, 45*, 340–363. https://doi.org/10.1177/0093854X17739928

Prochaska, J. O., & DiClemente, C. C. (1984). *The transtheoretical approach: Crossing traditional boundaries of therapy*. Homewood, IL: Dow Jones-Irwin.

Prochaska, J. O., DiClemente, C. C., & Norcross, J. C. (1992). In search of how people change: Applications to addictive behaviors. *American Psychologist, 47*, 1102–1114. https://doi.org/10.1037/0003-066X.47.9.1102

Rice, M. E., & Harris, G. T. (2005). Comparing effect sizes in follow-up studies: ROC area, Cohen’s d, and r. *Law and Human Behavior, 29*, 615–620. https://doi.org/10.1007/s10979-005-6832-7

Serin, R. C. (2007). *The Dynamic Assessment for Offender Re-entry (DRAOR) User Manual*. Ottawa, Ontario, Canada: Carleton University.

Serin, R. C., Gobeil, R., & Preston, D. L. (2009). Evaluation of the persistently violent offender treatment program. *International Journal of Offender Therapy and Comparative Criminology, 53*, 57–73. https://doi.org/10.1177/0306624X071313985

Serin, R. C., Lloyd, C. D., Helmus, L., Derkzen, D. M., & Luong, D. (2013). Does intra-individual change predict offender recidivism? Searching for the Holy Grail in assessing offender change. *Aggression and Violent Behavior, 18*, 32–53. https://doi.org/10.1016/j.avb.2012.09.002

Spielberger, C. D., Jacobs, G., Russell, S. F., & Crane, R. S. (1983). Assessment of anger: The State-Trait Anger Scale. In J. N. Butcher & C. D. Spielberger (Eds.), *Advances in personality assessment*, vol. 2 (pp. 159–187). Hillsdale, NJ: Lawrence Erlbaum.

Stefanska, E. B., Higgs, T., Carter, A. J., & Beech, A. R. (2018). “We boil at different degrees”: Factors associated with severity of attack in sexual killing. *Journal of Interpersonal Violence*. Advance online publication. https://doi.org/10.1177/0886262018758333

Suter, J. M., Byrne, M. K., Byrne, S., Howells, K., & Day, A. (2002). Anger in prisoners: Women are different from men. *Personality and Individual Differences, 32*, 1087–1100. https://doi.org/10.1016/S0191-8869(0)00105-2

Tew, J., Harkins, L., & Dixon, L. (2013). What works in reducing violent re-offending in psychopathic offenders. In L. A. Craig, L. Dixon & T. A. Gannon (Eds.), *What works in offender rehabilitation: An evidence-based approach to assessment and treatment* (pp. 129–141). Chichester, UK: John Wiley & Sons.
Wakeling, H. C., Beech, A. R., & Freemantle, N. (2013). Investigating treatment change and its relationship to recidivism in a sample of 3773 sex offenders in the UK. *Psychology, Crime & Law, 19*, 233–252. https://doi.org/10.1080/1068316X.2011.626413

Ward, T. (2016). Dynamic risk factors: Scientific kinds or predictive constructs. *Psychology, Crime & Law, 22*(1–2), 2–16. https://doi.org/10.1080/1068316X.2015.1109094

Williams, T. Y., Boyd, J. C., Cascardi, M. A., & Poythress, N. (1996). Factor structure and convergent validity of the aggression questionnaire in an offender population. *Psychological Assessment, 8*, 398–403. https://doi.org/10.1037/1040-3590.8.4.398

Wong, S. C. P., & Gordon, A. (1999–2003). *Violence Risk Scale*. Saskatchewan, CA: University of Saskatchewan.

Wong, S. C. P., & Gordon, A. (2006). The validity and reliability of the Violence Risk Scale: A treatment-friendly violence risk assessment tool. *Psychology, Public Policy, and Law, 12*, 279–309. https://doi.org/10.1037/1076-8971.12.3.279

*Received 20 March 2019; revised version received 3 April 2020*