How Does Offender Rehabilitation Actually Work? Exploring Mechanisms of Change in High-Risk Treated Parolees

J. A. Yesberg and D. L. L. Polaschek
Victoria University of Wellington

Acknowledgements

We are grateful to the New Zealand Department of Corrections for supporting the research and access to data, the Victoria University of Wellington Scholarships Committee for the first author’s Doctoral Scholarship, and to Professor Garth Fletcher for his help with the analysis for this paper.

Authors

Julia A. Yesberg, PhD, Research Fellow, Jill Dando Institute of Security and Crime Science, University College London, 35 Tavistock Square, London WC1H 9EZ, United Kingdom, +44 (0)7490193458, J.Yesberg@ucl.ac.uk

Devon L. L. Polaschek, PhD, DipClinPsych, Professor, School of Psychology, & New Zealand Institute of Security and Crime Science, The University of Waikato, Private Bag 3105, Hamilton 3283, New Zealand, Devon.polaschek@waikato.ac.nz
Abstract

Offender rehabilitation is typically thought to have been successful if a higher proportion of a sample of treatment completers avoids being reconvicted for an offence than a comparison sample. Yet this type of evaluation design tells us little about what brings about these outcomes. In this study we test whether change in dynamic risk factors during treatment is a recidivism-reducing mechanism in a sample of high-risk offenders. We also examine the extent to which change after treatment – in the period of re-entry from prison to the community – mediates this relationship. We found that although individuals made statistically significant change during treatment, this change was not significantly related to recidivism. We did, however, find tentative support for an indirect relationship between treatment change and recidivism, through change that occurred during re-entry. These findings signal the importance of the re-entry period for understanding how change in treatment is related to longer-term outcomes.
How Does Offender Rehabilitation Actually Work? Exploring Mechanisms of Change in High-Risk Treated Parolees

How do we know whether treatment has worked? Psychological treatment with offenders is typically said to have “worked” if a treatment completer avoids being convicted of a new offence in a specified follow-up period. Conversely, treatment is said to have failed if the individual is reconvicted. Based on this criterion for success, research has shown offender rehabilitation can work: it can lead to modest reductions in recidivism (Bonta & Andrews, 2017). But whether a treatment completer somehow manages to avoid a reconviction following treatment tells us little about how offender rehabilitation works. This study explores possible mechanisms of change in a sample of high-risk male offenders who completed intensive prison-based rehabilitation and were released from prison onto parole.

Figure 1 depicts a typical treatment outcome evaluation. Individuals convicted of an offence enter prison and serve part of their sentence before being offered a place in a treatment programme. Depending on the type of treatment, those who complete it will spend a number of months in the programme and most will subsequently re-enter the community. Some time later the evaluator follows up on reoffending outcomes; if the rate of recidivism for treatment completers is lower than a comparison group, treatment is said to be effective.

[Figure 1 near here]

Although not explicitly tested in typical outcome evaluations, the presumed mechanism that underlies successful treatment is the reduction of dynamic factors that cause and maintain criminal behaviour (e.g., criminal attitudes, poor emotional control; need principle; Bonta & Andrews, 2017). What is also assumed but not tested in typical outcome evaluations is that whatever change was made on these factors during treatment parallels what happens following treatment. In this study we test these assumptions (see Figure 2). First, we explore whether change in dynamic risk factors during treatment is a mechanism that leads to
Mechanisms of change in offender rehabilitation

reductions in recidivism. Then, we examine the extent to which change after treatment mediates this relationship.

[Figure 2 near here]

**Change in Dynamic Risk Factors as a Recidivism-Reducing Mechanism**

According to prominent models of offender rehabilitation, change in dynamic risk factors is considered to be the central mechanism that explains how treatment works (Bonta & Andrews, 2017). Yet little research has tested this mechanism empirically (Monahan & Skeem, 2014). Interventions that target dynamic risk factors have been shown to significantly reduce recidivism (Bonta & Andrews, 2017). Those that target multiple dynamic risk factors have shown to bring about the greatest reductions (Gendreau, French, & Taylor, 2002). These findings are usually interpreted as evidence that change in dynamic risk factors is a recidivism-reducing mechanism. But, in fact, there are three types of evidence needed to confirm this assumption (Kroner & Yessine; 2013). First, as above, the treatment needs to be shown to reduce recidivism. Second, participants need to demonstrate change in the factors identified and targeted during treatment. And third, the amount of change an individual makes needs to be related to their likelihood of reoffending.

Typically, participants are assessed prior to treatment on a range of areas related to their offending and then re-assessed on the same variables following treatment, to determine progress. Treatment change is most often assessed empirically using offender self-report (Polaschek, Bell, Calvert, & Takarangi, 2010) or staff rated instruments designed specifically to measure change in dynamic risk factors over time (e.g., third- and fourth-generation instruments; Douglas & Skeem, 2005). Research using both types of instruments has found clinically and statistically significant pre- to post-treatment changes, in areas such as antisocial attitudes, cognitive distortions, anger expression, and impulsivity (e.g., Beggs & Grace, 2011; Hudson, Wales, Bakker, & Ward, 2002; Kingston & Olver, 2018; Lewis, Olver,
Mechanisms of change in offender rehabilitation

& Wong, 2013; Polaschek & Dixon, 2001; Woessner & Schwedler, 2014). These findings suggest that treatment can alter dynamic risk factors. However, to determine whether this treatment change is a recidivism-risk reducing mechanism, research needs to demonstrate that altering these factors leads to reductions in recidivism (Kroner & Yessine, 2013).

There is surprisingly little research linking treatment change and recidivism. In Serin and colleagues’ (2013) review of the literature, only 17 out of 378 studies measuring treatment change attempted to link change to recidivism, and just 10 showed a significant association between recidivism and at least one change measure. Most of the studies reviewed by Serin were based on self-report measures, suggesting an offender’s own evaluations of change may not be externally valid. Furthermore, many of the studies did not control for baseline risk. Previous research has found those who are highest risk prior to treatment have the greatest capacity to change (Beggs & Grace, 2011; Olver, Kingston, Nicholaichuk, & Wong, 2014), so the lack of a link between treatment change and recidivism could simply reflect the higher risk of the people who change the most in treatment. However, recent studies not included in Serin’s review that do control for pre-treatment risk—either by adding it as a covariate in the analyses or creating residualised change scores—have found little or no relationship between self-reported change and reductions in recidivism (Barnett, Wakeling, Mandeville-Norden, & Rakestrow, 2013; Kroner & Yessine, 2013; Olver et al., 2014; Woessner & Schwedler, 2014; see Beggs & Grace, 2011; Hudson et al., 2002 for exceptions).

Research examining change in externally rated measures shows similarly inconclusive results, although some findings from this line of research are promising. In a study of 150 violent offenders with significant psychopathic personality traits, after controlling for pre-treatment levels of risk, men who made more change on the Violence Risk Scale (VRS; Wong & Gordon, 2000) during treatment were significantly less likely to reoffend on release than those who made less change (Lewis et al., 2013). Similar results have been found with
Mechanisms of change in offender rehabilitation

the Violence Risk Scale: Sexual Offender Version, with both Canadian and New Zealand samples (Olver, Christofferson, Grace, & Wong, 2013; Olver, Wong, Nicholaichuk, & Gordon, 2007; Olver, Kingston, Nicholaichuk, & Wong, 2014). Furthermore, in a study of forensic psychiatric patients changes in risk and protective factors were significantly related to violent recidivism at 1- and 11-year follow-ups (de Vries Robbé, de Vogel, Douglas, & Nijman, 2015). Most of these studies are retrospective in design and rely on ratings made by researchers external to the treatment process (see Olver et al., 2014 for an exception). A number of studies using change ratings made by the therapists who provided the treatment have found little relationship between within-treatment change and recidivism (Barbaree, 2005; Seager, Jellicoe, & Dhaliwal, 2004; Quinsey, Khanna, & Malcolm, 1998). However, other studies using repeated measurements of dynamic risk in routine practice (e.g. during probation supervision) have found some links between changes in dynamic risk and recidivism (Cohen, Lowenkamp, & VanBenschoten, 2016; Howard & Dixon, 2013; Labrecque, Smith, Lovins, & Latessa, 2014; Vose, Smith, & Cullen, 2013).

**Why is Treatment Change Rarely Related to Recidivism?**

The most commonly cited explanations for the lack of association between treatment change and recidivism have to do with the measurement of treatment change (e.g., predictive validity of change measures; Mills, Loza, & Kroner, 2003; Polaschek, et al., 2010), but there are a number of conceptual explanations often overlooked. One explanation relates to the way change is conceptualised in these studies and the lack of recognition of the process by which people change their behaviour. For example, despite measuring dynamic and malleable factors, in most treatment change studies, change is considered to be over at the end of treatment (Quinsey, Jones, Book, & Barr, 2006). Yet we know from theories of behaviour change (e.g., the Transtheoretical Model of Change; TTM; Prochaska, DiClemente, & Norcross, 1992) that change is not a discrete event but a process of qualitatively distinct
Mechanisms of change in offender rehabilitation

stages through which people progress. This process often involves periods of relapse back to earlier stages before permanent change is achieved (Casey, Day, & Howells, 2005).

Research that attempts to link within-treatment change to recidivism makes the assumption that (1) there is no additional change in the assessed variables following treatment or, (2) if there is, that such post-treatment change parallels the change made in treatment. However, recent research suggests behaviour change with offenders does not necessarily conform to these assumptions. In a small study of 35 life-sentenced prisoners who completed intensive treatment, we found that the direction and volume of within-treatment change on the VRS did not necessarily parallel change made in the 6-12 months after treatment completion (Yesberg & Polaschek, 2014). Some men who made significant change during treatment increased their risk in the follow-up period, while another group continued to decrease their risk and a third group made no further change. Similarly, one study found that all offenders changed in a positive direction during treatment, but after release, the scores of men who were reconvicted had returned to approximately pre-treatment levels, whereas men who were not reconvicted continued to improve (Polaschek & Dixon, 2001).

Theories of desistance from crime are also relevant here. These theories suggest that desistance does not occur instantaneously, and often involves a number of false starts (Walker, Bowen, Brown, & Sleath, 2014). Desistance is thought of as a gradual process involving the interaction of multiple factors – both internal and external to the individual – that influence the transition out of crime (LeBel, Burnett, Maruna, & Bushway, 2008; Serin & Lloyd, 2009). Göbbels, Ward and Willis’ (2012) Integrated Theory of Desistance from Sexual Offending (ITDSO) suggests there are four phases in the desistance process: (1) decisive momentum (initial desistance), (2) rehabilitation (promoting desistance), (3) re-entry (maintaining desistance), and (4) normalcy (maintaining desistance over a long period of time). According to the ITDSO, during the rehabilitation phase, offenders reconstruct their
identities, acquire strategies to successfully live a pro-social life, and overcome internal and external obstacles (e.g., dynamic risk factors). During the re-entry phase, there a number of factors that act as facilitators (e.g., high expectations, positive social capital) and barriers (e.g., lack of employment, housing) to re-entry. Göbbels and colleagues (2014) write that a “successful re-entry phase facilitates the ex-offender’s achievement of long-term desistance” (p. 356). The final phase of the ITDSO is an extension of the re-entry phase where an offender has maintained his commitment to change and has desisted from crime for a long period of time. Taken together, this research suggests desistance is dependent on more than whether or not an offender changes during treatment: they also need to maintain a commitment to change and manage a number of barriers and facilitators during re-entry.

We know the period of re-entry from prison back into the community is a particularly challenging time and recidivism rates are at their peak (Burnett, 2009); in New Zealand, about a quarter of all high-risk parolees released into the community return to prison in the first 100 days (Nadesu, 2007). Yet we know little about how treatment change is related to what goes on during this time, and how this may affect longer-term recidivism rates. Thus, taking into account what happens during the re-entry process should increase our understanding of the mechanisms by which rehabilitation leads to long-term reductions in recidivism and the achievement of desistance.

**Introduction to the Current Study**

Although typical outcome evaluations can tell us whether the treatment reduced subsequent rates of recidivism, they provide little information about *how*. This study tests two assumptions underlying successful treatment with a sample of high-risk offenders released onto parole in New Zealand. We have already found with this sample that treated offenders were significantly less likely to be reconvicted than men who had not undertaken treatment (Polaschek, Yesberg, Bell, Dickson & Casey, 2016). Here, we extend that research in two
Mechanisms of change in offender rehabilitation

ways. First, we examine whether change in dynamic risk during treatment is a mechanism that leads to reductions in recidivism. We use pre- and post-treatment VRS ratings and two indices of recidivism: violent reconviction and reimprisonment. Prior research has found a significant relationship between change on the VRS and recidivism (Lewis et al., 2013), but this research relied on ratings made retrospectively by researchers. One study using the VRS-SO found clinician-rated change to be predictive (Olver et al., 2014), but no equivalent study has been conducted with the VRS. Therefore, we address a gap in the literature by using VRS ratings made by the clinicians who provided the treatment (i.e., field ratings).

Then, we explore whether change during the period of re-entry from prison to the community mediates the relationship between treatment change and recidivism. We use the Dynamic Risk Assessment for Offender Re-entry (DRAOR; Serin, Mailloux, & Wilson, 2012)—a measure of stable and acute dynamic risk factors and protective factors used by probation officers in the community—to explore this relationship. The latter part of this study is particularly novel; no empirical research that we know of has examined how treatment change is related to re-entry change.

Method

Sample

The sample comprised 123 men who completed an intensive treatment programme housed in one of four purpose built units throughout New Zealand’s national prison system—the High-Risk Special Treatment Units (HRSTUs; see Polaschek & Kilgour, 2013 for a full description of the programme)—and were subsequently released into the community on parole between December 2010 and November 2013. II The HRSTUs aim to treat about 120 high-risk violent male prisoners each year. Prisoners eligible for admission into one of the HRSTUs usually have an estimated risk of returning to prison of at least 70% over the five years following release, are serving imprisonment sentences of at least 2 years, are over the
age of 20, have a low-medium or minimum security rating, have sufficient time left on their sentence to complete the programme, and agree to be transferred to one of the units.

**The High-Risk Special Treatment Units (HRSTUs)**

The treatment model of the HRSTUs provides a structured, closed-group cognitive-behavioural therapy intervention within a modified democratic therapeutic community. The core treatment programme is delivered to groups of 10 men by pairs of facilitators. Men attend group sessions for approximately 250 hours over 25 weeks, and remain in the treatment unit for 10-12 months. Prisoners may also receive individual psychological treatment for problems that make it hard for them to fully participate in group sessions (e.g., social anxiety) and additional interventions may be provided for Māori that focus on specific cultural needs (Polaschek & Kilgour, 2013). The group sessions focus on a variety of areas including offence-supportive thinking, mood management, problem solving, and relationships/communication (Kilgour & Polaschek, 2012). The final part of treatment focuses on preparing men for release. In addition to basic release planning (i.e., accommodation, support networks), offenders develop a personalised safety plan where they identify potential high-risk situations and develop strategies to effectively manage them.

Outcome evaluations from the HRSTUs are limited but positive overall. The most recent evaluation using a quasi-experimental design found a significant reduction in recidivism within one year after release for HRSTU completers (including the men in the current study) compared with an equivalent comparison group (Polaschek et al., 2015). Effect sizes ranged from weak ($\Phi = .15$) for violent reconviction to moderate for breach of parole, any new criminal conviction, and reimprisonment ($\Phi = .21, .21, \text{ and } .24$, respectively). These effect sizes are larger than earlier evaluations of the HRSTUs (e.g., Kilgour & Polaschek, 2012; Polaschek, 2011) and would usually be taken to indicate a noteworthy treatment effect (e.g., Bonta & Andrews, 2017).
Sample Characteristics

The 123 HRSTU completers in this sample had a mean 76\% likelihood of returning to prison in the five years following release.\textsuperscript{iii} The majority identified as Māori (63\%), 31\% as New Zealand European, 4\% as Pasifika, and the remainder identified as Other European. Sample members ranged in age from 19 to 56 years at the time they were released from prison ($M = 32$ years, $SD = 8$ years). They had an average of 71 previous convictions ($SD = 54$), including 5 for violence ($SD = 4.1$). They were aged 16 on average at the date of their first conviction ($SD = 2.1$) and 19 at the date of their first violent conviction ($SD = 3.0$). The index offence for almost two-thirds of the sample was a violent offence.

Nine men were sentenced to life imprisonment for their index offence(s).\textsuperscript{iv} The remaining 114 men had determinate sentence lengths, which were 4 years on average ($SD = 2.3$). Although the programme was designed for prisoners to complete just prior to release, in reality, many HRSTU completers spend a number of months in prison after the programme, often being transferred to mainstream units throughout the national prison system as they prepare for release. The average number of days from treatment completion to release was 210 ($SD = 232$). At the time of release, they had served an average of 4.3 years in prison ($SD = 4.5$) and 79\% were released before their sentence end date (i.e., were granted early parole), while 21\% were released at the end of their sentence. The average length of parole they were to complete was 391 days ($SD = 217$).

Measures

The Violence Risk Scale. The Violence Risk Scale (VRS; Wong & Gordon, 2000) is a 26-item staff-rated risk instrument that assesses 6 static (e.g., age at first violent offence) and 20 dynamic (e.g., criminal attitudes, impulsivity, interpersonal aggression) risk factors, particularly designed to measure change in custodial treatment. Each item is rated prior to treatment on a 4-point scale from completely unrelated to violence (0) to strongly related to
violence (3). The dynamic items considered problematic (those rated either a 2 or a 3) receive a second rating based on the offender’s current level of engagement in change (‘stage of change’) on that item (i.e., pre-contemplation, contemplation, preparation, action, or maintenance) to determine the degree to which the offender currently acknowledges the risk factor as problematic and has made observable efforts to overcome it (Wong & Gordon, 2000, 2009).

Post-treatment scores are obtained by re-rating stage of change for all dynamic risk factors originally rated 2 or 3, and subtracting 0.5 points from the pre-treatment score for progression through each stage of behaviour change. For this purpose, pre-contemplation and contemplation are considered equivalent; risk scores do not reduce if the only change on the risk factor is a shift from not acknowledging the factor (pre-contemplation) to acknowledging it (contemplation). For example, if an offender who received a rating of ‘2’ for the dynamic risk factor ‘criminal attitudes’ moved from a pre-treatment stage of change rating of ‘pre-contemplation’ to a post-treatment stage of change rating of ‘contemplation’, the final risk score for this item would remain as 2. However, if the offender moved to a post-treatment rating of ‘preparation’, 0.5 points would be subtracted from the original score, leading to a final risk score of 1.5 for this item. If the offender moved to ‘Action’ on this item, the final risk score would reduce to 1 (Wong & Gordon, 2000). Thus, the VRS scores reflect changes in dynamic risk over time (i.e., movement from problem recognition, to initiating behaviour change, to increasing resilience of change). Previous research from New Zealand has found VRS scores to be significantly predictive of recidivism (e.g., AUC = .73; Dickson, Polaschek, & Casey, 2013).

**DRAOR.** The Dynamic Risk Assessment for Offender Re-entry (DRAOR; Serin et al., 2012) is, like the VRS, a dynamic tool, but in contrast is designed to facilitate the assessment of recidivism risk in the community and to inform case planning and risk
Mechanisms of change in offender rehabilitation

management. The instrument comprises 19 items, divided into 3 subscales: stable dynamic risk factors, acute dynamic risk factors, and protective factors. Each item is rated using a three-point scoring format (0, 1, 2). A score of ‘0’ indicates the absence of the item, a ‘2’ indicates it is strongly present, and a ‘1’ rating is used to indicate it is somewhat present, or the evidence is inconsistent. The items were derived from a review of the literature on violent offender risk assessment and on desistance, and theoretically organised into three subscales.

We explored the subscale structure in a previous study with a sample of high-risk offenders (Yesberg 2015; Yesberg & Polaschek, 2015). Findings revealed four components rather than the theoretically derived three-subscale structure. The Protective subscale was identical in the new structure but the majority of the acute risk factors split to form two subscales (Internal and External Acute). One original acute item loaded on the Stable subscale and one original Stable item loaded on the new External Acute subscale. This four-subscale DRAOR is presented in the figure below and will be used in this study. We also calculated a DRAOR total score by summing the stable and acute risk items and subtracting the protective items. Total scores range from -12 to 26; higher scores indicate greater recidivism risk.

[Figure 3 near here]

In New Zealand, all offenders released from prison onto parole are scored on the DRAOR multiple times during their sentence. Supervising probation officers score the DRAOR during every report-in or non-trivial contact they have with the offender. Depending on an offender’s initial risk level and how long they have been on parole, the DRAOR could be administered twice weekly to fortnightly, or even at longer intervals. To score the DRAOR, probation officers use information gathered from interviews with offenders, their families or partners, treatment providers, and other external sources (e.g., police intelligence activity). DRAOR scores have been found to be predictive of recidivism in high-risk offenders (Yesberg & Polaschek, 2015), female offenders (Yesberg, Scanlan, & Polaschek,
Mechanisms of change in offender rehabilitation

2015), youth (Fortune et al., 2015) and a large sample of New Zealand parolees \(N = 3498;\) Davies, 2019; Hanby, 2013; Lloyd, 2015). Originally developed in Canada, the DRAOR has recently been implemented in other jurisdictions (e.g., Iowa; Chadwick, 2014).

**Recidivism.** Recidivism data were extracted from New Zealand’s national conviction records database between October 2013 and September 2014. Two indices of recidivism were examined: violent reconviction and any reconviction leading to reimprisonment.³ We set recidivism to 1 year after release. In that time, 14% of men were convicted for violence and 30% received a sentence of imprisonment.

**Procedure**

**VRS Ratings.** Pre- and post-treatment VRSs rated by therapists as part of routine clinical practice were collected. In the HRSTUs, a member of the therapy staff who has been trained in the use of the VRS completes the pre- and post-treatment ratings using file information, staff observational data, and information gathered through interviews with the offender. To ensure the VRS was up to date at release, in addition to the therapist-rated VRSs, trained research assistants completed a third VRS for men who remained in prison for more than 6 months after the programme \(n = 31\). The ratings were based on information from interviews with the offender just prior to release and relevant file data (e.g., psychological reports, assessment reports prepared for Parole Board appearances, misconduct and incident reports, results from drug tests). The rater reviewed the pre- and post-treatment ratings completed by therapists and re-rated stage of change for all dynamic items considered to be treatment targets (rated a 2 or 3 at pre-treatment) using new information collated from the date men’s post-treatment VRS was completed until the date they were released from prison. Therefore, all sample members had a VRS based on pre-release information and either one or
two additional VRSs completed by therapists and research assistants as they progressed through their prison sentence to release.

No inter-rater reliability data were available for the VRSs scored in therapy; they are subjected to scrutiny from clinical supervisors during their completion. The VRSs completed by the research assistants were subject to inter-rater reliability as part of a larger research project. Overall, inter-rater reliability was “almost perfect” for the static items (Landis & Koch, 1977): $\kappa = 0.97, p<.001$, and very good for the dynamic items: $\kappa = 0.89, p<.001$, and stage of change: $\kappa = 0.88, p<.001$. Item Kappas ranged between 0.63 and 1.00.

**DRAOR Ratings.** DRAOR scores completed by probation officers in routine parole practice were extracted from electronic offender records. Because the DRAOR was used in this study to measure change during re-entry, we isolated for analysis the first and last scores within the first 100 days of release. On average, the first rating was made 1.7 days after release ($SD = 3.2$) and the last rating 87 days after release ($SD = 18.8$). We calculated net change scores by subtracting the last rating within 100 days from the first rating following release. No inter-rater reliability data were available for the DRAOR.

**Results**

First, we test whether change on the VRS during treatment is related to a reduction in recidivism. Recall that previous research has found that the HRSTUs significantly reduce rates of recidivism (Polaschek et al., 2015). The two additional findings needed to determine whether treatment change is a recidivism-reducing mechanism are: (1) that offenders change during treatment and (2) that change is related to recidivism. The second part of the results explores the additional influence of change during the re-entry period.

**How Much Change Occurs During Treatment?**

Pre- and post-treatment VRS scores are presented in Table 1, along with change scores. Using Wong and Gordon’s (2006) recommendation of a VRS total score of 50 as a
cut-off to indicate a high risk group, 99 men (79%) were considered at high risk of violent reoffending, prior to treatment. The mean pre-treatment VRS dynamic scale score of 42.4 ($SD = 7.0$) indicates that, on average, each dynamic item was given a rating of 2.1, identifying that item as sufficiently strongly related to violence to be a worthwhile treatment target. The most common treatment targets (items rated 2 or 3) were violent lifestyle, criminal attitudes, criminal peers, released to high-risk situations, violence cycle, and cognitive distortion. Stage of change scores, in addition to providing anchors for measuring progress, indicate current level of engagement in change. Prior to beginning treatment, for those items rated 2 or 3, men were at the Contemplation stage of change on average. A rating of Contemplation indicates the offender acknowledges the problem area but has made no behavioural steps to address it.

Post-treatment dynamic scale scores were significantly lower than the pre-treatment scores, $t(122) = 19.27, p < .001, d = 0.69, 95\% CI = -0.55, 1.92$, suggesting that men made statistically significant change during treatment. Treatment change scores were computed by subtracting the post-treatment VRS dynamic score from the pre-treatment VRS dynamic score, based on the method outlined in the VRS manual (Wong & Gordon, 2000). As Table 1 shows, men in the sample had changed an average of 4.6 points (range 0 to 13.5), which is equivalent to a single-stage shift on 9 dynamic risk factors. At the end of treatment, men’s average stage of change per treatment target was nearing Preparation. There was a small but significant correlation between men’s pre-treatment dynamic score and the amount of change they made ($r = .22, p = .017$), indicating that the higher their dynamic risk score prior to treatment, the more change they made during the programme.

[Table 1 near here]

**Is Treatment Change Related to Recidivism?**

During the first year in the community, 14% of offenders were convicted for a new violent offence and 30% received a new sentence of imprisonment. Logistic regression was
used to examine the relationship between treatment change and recidivism, statistically controlling for pre-treatment levels of static and dynamic risk (VRS). The dependent variable was recidivism (violent conviction or reimprisonment within 1 year). The pre-treatment VRS score (static or dynamic) was entered as a covariate in the first block followed by the treatment change score in the second block. The VRS static score predicted reimprisonment, \( \chi^2(1, N = 123) = 10.38, \ p = .001 \), but not reconviction for violence: \( \chi^2(1, N = 123) = 3.19, \ p = .074 \). Conversely, the VRS dynamic score predicted neither reconviction for violence: \( \chi^2(1, N = 123) = 1.93, \ p = .165 \), nor reimprisonment, \( \chi^2(1, N = 123) = .37, \ p = .542 \). As shown in Table 2, treatment change did not significantly predict either type of recidivism outcome when added into the models, suggesting that after controlling for pre-treatment levels of risk, the amount of change an offender made in treatment was not significantly related to his likelihood of recidivism. Note, however, that the relationships were in the predicted direction (i.e., regression coefficient values were negative, suggesting more change was associated with lower rates of recidivism).

[Table 2 near here]

**Change During Re-entry as a Mediator**

Next, using the DRAOR, we explored the additional influence of change during re-entry. In other words, we asked whether change that occurs after treatment—and in this case, after an offender has re-entered the community—can explain the lack of significant relationship between treatment change and recidivism. Recall that DRAOR change scores were calculated by subtracting from the first score, the last score within 100 days of release. Means and standard deviations for the first and last DRAOR scores (and net change scores) are presented in Table 3; higher scores on the three risk subscales indicate greater risk whereas higher scores on the protective subscale indicate greater protection. As shown in the
table, all subscales changed during re-entry; risk scores decreased and protective scores increased. Change on all subscales was significant except for the Stable subscale.

To explore whether change during re-entry mediated the relationship between treatment change and recidivism, regression models were tested using the procedure outlined by Hayes (2017). In all regressions, pre-treatment VRS dynamic scores and initial DRAOR scores were controlled for. Men were excluded if their first reoffence leading to reconviction occurred in the first 100 days after release and before their last DRAOR score. For violent reconviction, the sample size dropped to 120 because 3 men reoffended violently before the date of their last DRAOR score; for reimprisonment, the sample size dropped to 112 because the date of the offence that led to imprisonment was before the last DRAOR score for 11 men.

Regression coefficients and significance values are presented in Table 4. The first regression models (Ordinary Least Squares) tested the relationship between the predictor variable (X) and the potential mediators (M). Treatment change significantly predicted change on the Stable, Internal Acute and Protective subscales, and the DRAOR total score; it did not predict change on the External Acute subscale. These results suggest that men who made more change in treatment went on to make more change during re-entry (i.e., reduced risk and increased protective factors). The second regression models (logistic) included the predictor variable (X) and potential mediators (M) as predictors of recidivism (Y). Only one potential mediator was significantly independently predictive of recidivism: change on the Stable subscale predicted reimprisonment (see Figure 4). The bootstrap confidence intervals derived from 500 samples indicated that the indirect effect coefficient was statistically significant, b = -.06, SE = .03, 95% CI = -.13, -.01.

Discussion
Understanding how treatment works is crucial for designing and delivering the most effective programmes. Research evaluating offender rehabilitation has established that treatment can reduce recidivism, but less research has directly explored what brings about these reductions. This study tested two assumptions underlying effective offender rehabilitation. First, we examined whether change in dynamic risk factors is a mechanism that brings about reductions in recidivism (need principle; Bonta & Andrews, 2017). We found that although offenders made statistically significant change on the VRS during treatment, the amount of change they made was not significantly related to their subsequent rates of recidivism, even after controlling for pre-treatment levels of static and dynamic risk.

This finding is mainly in line with the empirical literature as a whole, but is at odds with research using the same instrument. Studies using both the VRS and the VRS-SO have found significant associations between within-treatment change and subsequent rates of violent and sexual recidivism, respectively (Lewis et al., 2013; Olver et al., 2014; Olver et al., 2007). The findings with the VRS-SO have also been replicated with a sample of child sex offenders who completed a sex offender treatment programme in New Zealand, similar to the HRSTUs (Beggs & Grace, 2011; Olver et al., 2013). There are a couple of key differences between the current study and the other VRS change studies. First, this study used ratings made by clinicians in routine practice whereas most other change studies were retrospective in design: researchers made ratings from file information after treatment had been completed. Although clinician-rated measures have been shown to be valid predictors of recidivism, there may still be biases associated with using clinician ratings of change because of their close involvement with the offenders’ treatment (Seto, 2003). One study conducted prospectively using clinician-rated VRS-SO assessments did find a significant relationship between change and recidivism (Olver et al., 2014), which is at odds with the present findings.
Other differences between this study and the other VRS change studies include a smaller sample size, a shorter follow-up time, and lower base rates of recidivism (14% base rate of violent recidivism in this study compared to a 46% base rate in the Lewis et al. study): all variables that impact on a study’s ability to detect significant differences (i.e., statistical power). And, if we compare the effect sizes of the current study with the previous VRS change studies, they are actually remarkably similar ($e^B = .91-.92$ in Lewis et al. versus $.86-.93$ in the current study), suggesting the possibility that Type II errors are present.

Nonetheless, there are some conceptual issues with attempting to link treatment change with long-term recidivism outcomes. One issue is that we are effectively treating treatment change as a static predictor. As Quinsey et al. (2006) wrote, “although change in the prerelease period can be measured and related to postrelease outcome, this change is nevertheless over at the time the follow-up period begins” (p. 1540). We know from theories of behaviour change (e.g., the Transtheoretical Model, Prochaska et al., 1992) and the desistance literature (Göbbels et al., 2012) that change is a process of stages, and that people often experience relapse back to earlier stages of change before lasting change is achieved. Thus, the lack of a significant relationship between treatment change and recidivism could simply be due to the fact that we are missing change that occurs in the follow-up period. The second part of this study explored whether post-release change (i.e., change measured during re-entry) mediates the relationship between treatment change and recidivism; no research that we know of has asked this question before.

We found, firstly, that treatment change was related to change during re-entry (i.e., a decrease in risk factors and an increase in protective factors). Secondly, we found tentative support for an indirect relationship between treatment change and recidivism through change on dynamic risk factors during re-entry. Men who made more change in treatment made more change on stable dynamic risk during re-entry, which was in turn related to lower rates of
reimprisonment. Whether the change individuals made during re-entry was a continuation of the change made in treatment is unclear, because change across the two time intervals was measured with different instruments. However, the results certainly suggest that change is not over at the point of release and that exploring what happens during re-entry may lead to a more complete understanding of the process by which offenders change their behaviour and desist from crime.

One major limitation to this study is that we did not include a comparison group who were measured on the same variables over a similar period of time. As a result, it is unclear whether the programme caused the change in risk as opposed to some other variable, such as repeated assessment or regression to the mean (Monahan & Skeem, 2014). Future research should include a comparison group to determine if what we assume to be treatment effects is actually brought about by the programme. This study suffered from low statistical power, especially with regards to sample size and recidivism base rates. Furthermore, limiting our sample to high-risk offenders constrained the amount of variability in the predictor variables (VRS and DRAOR, both risk instruments), which could explain their poor predictive validity.

Future research should take a more sophisticated approach to measuring change, especially when using an instrument like the DRAOR, which does not include a theoretical model for measuring change. Subtracting one score from another does not provide any information regarding rates of change or fluctuations in change that may occur during a follow-up period. Although we used this approach in the current study to be comparable to the overall change measured during treatment, for an instrument like the DRAOR – which incorporates acute dynamic risk factors and is designed for multiple measurements – the net change score is a rather crude indicator of the progress an offender may have made during re-entry. Future research could adopt longitudinal designs with multiple assessment points, and use more sophisticated statistical techniques to examine rates of change over time, such as
growth curve analysis or multi-level modelling approaches (see Polaschek & Yesberg, 2018). Additional qualitative research should also be conducted to explore offenders’ experiences of the change process, and whether the re-entry period helps or hinders desistance.

In summary, findings from this study suggest that to better understand how treatment works to reduce recidivism and facilitate desistance, outcome evaluations should focus not only on the amount of change an individual makes in treatment but also the extent to which this change continues after release. The current study tentatively signals the importance of the re-entry period for understanding how change in treatment is related to longer-term outcomes. Further explorations into mechanisms of change in offender rehabilitation are needed to ensure we are designing and delivering the most effective programmes.
References


Mechanisms of change in offender rehabilitation


Table 1

*Descriptive Statistics for Pre- and Post-treatment VRS Scores and Change Scores*

<table>
<thead>
<tr>
<th>VRS measure</th>
<th>Pre-treatment</th>
<th></th>
<th>Post-treatment</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M (SD)</td>
<td>Range</td>
<td>M (SD)</td>
<td>Range</td>
</tr>
<tr>
<td>Static</td>
<td>12.9 (2.5)</td>
<td>6 to 17</td>
<td>12.8 (2.5)</td>
<td>6 to 17</td>
</tr>
<tr>
<td>Dynamic</td>
<td>42.4 (7.0)</td>
<td>21 to 53</td>
<td>37.6 (7.0)</td>
<td>18 to 52.5</td>
</tr>
<tr>
<td>Total</td>
<td>55.3 (7.8)</td>
<td>29 to 69</td>
<td>50.5 (7.9)</td>
<td>26.5 to 65.5</td>
</tr>
<tr>
<td>Stage of Change</td>
<td>2.0 (.48)</td>
<td>1.1 to 4</td>
<td>2.8 (.48)</td>
<td>1.3 to 4.1</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Treatment Change</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dynamic Total</td>
</tr>
</tbody>
</table>

*a For these analyses stage of change was coded as follows: 1=pre-contemplation, 2=contemplation, 3=preparation, 4=action, and 5=maintenance. For all dynamic items rated 2 or 3, the stage of change for those items was averaged. Note that this scale differs from the scale used to obtain change scores. In the latter, pre-contemplation and contemplation are scored as equivalent (Wong & Gordon, 2000).*
Table 2

Logistic Regressions of Treatment Change Predicting Recidivism Controlling for Pre-treatment Static and Dynamic Risk

<table>
<thead>
<tr>
<th></th>
<th>Violent Conviction</th>
<th>Reimprisonment</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>B (SE)</td>
<td>Wald</td>
</tr>
<tr>
<td>Block 1  VRS Static</td>
<td>.18 (.11)</td>
<td>2.66</td>
</tr>
<tr>
<td>Block 2  VRS Static</td>
<td>.19 (.11)</td>
<td>2.83</td>
</tr>
<tr>
<td>Treatment Change</td>
<td>-.14 (.11)</td>
<td>1.68</td>
</tr>
<tr>
<td>Block 1  VRS Dynamic</td>
<td>.06 (.04)</td>
<td>1.75</td>
</tr>
<tr>
<td>Block 2  VRS Dynamic</td>
<td>.07 (.04)</td>
<td>2.48</td>
</tr>
<tr>
<td>Treatment Change</td>
<td>-.16 (.10)</td>
<td>2.25</td>
</tr>
</tbody>
</table>

**p < .01 Note. Significant values are highlighted in bold.**
### Table 3

*Means and Standard Deviations for DRAOR Subscale Total and Average Item Scores*

<table>
<thead>
<tr>
<th></th>
<th>Initial</th>
<th>Last</th>
<th>Change</th>
<th>$t$ (121)</th>
<th>$p$</th>
<th>$d$</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stable</td>
<td>6.80(2.24)</td>
<td>6.66(2.60)</td>
<td>.14(1.75)</td>
<td>.88</td>
<td>.382</td>
<td>.06</td>
<td>-.34, .52</td>
</tr>
<tr>
<td>Internal Acute</td>
<td>1.67(1.34)</td>
<td>1.10(1.26)</td>
<td>.58(1.60)</td>
<td>4.02</td>
<td>&lt;.001</td>
<td>.44</td>
<td>.20, .66</td>
</tr>
<tr>
<td>External Acute</td>
<td>4.16(1.23)</td>
<td>3.76(1.35)</td>
<td>.41(1.09)</td>
<td>4.16</td>
<td>&lt;.001</td>
<td>.31</td>
<td>.09, .55</td>
</tr>
<tr>
<td>Protective</td>
<td>5.91(1.75)</td>
<td>6.43(2.20)</td>
<td>-.52(1.73)</td>
<td>-3.36</td>
<td>.001</td>
<td>-.26</td>
<td>-.57, .13</td>
</tr>
<tr>
<td>DRAOR Total</td>
<td>6.70(4.58)</td>
<td>5.08(5.66)</td>
<td>1.57(4.26)</td>
<td>4.06</td>
<td>&lt;.001</td>
<td>.32</td>
<td>-.49, 1.32</td>
</tr>
</tbody>
</table>
### Table 4

**Regression Coefficients for Mediation Analyses of Indirect Effects**

<table>
<thead>
<tr>
<th></th>
<th>Stable Change (M)</th>
<th>Violent conviction (Y)</th>
<th>Reimprisonment (Y)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Coeff.</td>
<td>SE</td>
<td>p</td>
</tr>
<tr>
<td>Treatment Change (X)</td>
<td>$a$</td>
<td>.18</td>
<td>.06</td>
</tr>
<tr>
<td>Stable Change (M)</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>constant</td>
<td>$i_m$</td>
<td>.33</td>
<td>.96</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Stable Change (M)</th>
<th>Violent conviction (Y)</th>
<th>Reimprisonment (Y)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Coeff.</td>
<td>SE</td>
<td>p</td>
</tr>
<tr>
<td>Treatment Change (X)</td>
<td>$a$</td>
<td>.11</td>
<td>.04</td>
</tr>
<tr>
<td>Internal Acute Change (M)</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>constant</td>
<td>$i_m$</td>
<td>-.32</td>
<td>.68</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Stable Change (M)</th>
<th>Violent conviction (Y)</th>
<th>Reimprisonment (Y)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Coeff.</td>
<td>SE</td>
<td>p</td>
</tr>
<tr>
<td>Treatment Change (X)</td>
<td>$a$</td>
<td>.06</td>
<td>.04</td>
</tr>
<tr>
<td>External Acute Change (M)</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>
Mechanisms of change in offender rehabilitation

<table>
<thead>
<tr>
<th></th>
<th>Protective Change (M)</th>
<th>Violent conviction (Y)</th>
<th>Reimprisonment (Y)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Coeff.</td>
<td>SE</td>
<td>p</td>
</tr>
<tr>
<td>Treatment Change (X)</td>
<td>$a$</td>
<td>.15</td>
<td>.06</td>
</tr>
<tr>
<td>Protective Change (M)</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>constant</td>
<td>$i_m$</td>
<td>-2.51</td>
<td>1.22</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Protective Change (M)</th>
<th>Violent conviction (Y)</th>
<th>Reimprisonment (Y)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Coeff.</td>
<td>SE</td>
<td>p</td>
</tr>
<tr>
<td>Treatment Change (X)</td>
<td>$a$</td>
<td>.49</td>
<td>.14</td>
</tr>
<tr>
<td>DRAOR Total Change (M)</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>constant</td>
<td>$i_m$</td>
<td>3.16</td>
<td>2.31</td>
</tr>
</tbody>
</table>

*Note.* The pre-treatment VRS and initial DRAOR scores were both controlled for in these analyses. Significant values are highlighted in bold.
Figure 1

*Typical Treatment Outcome Evaluation*

![Typical Treatment Outcome Evaluation Diagram](image)

Figure 2

*Model for Exploring Mechanisms of Change*

![Model for Exploring Mechanisms of Change Diagram](image)
Figure 3

*Empirically Derived Four-Subscale DRAOR Structure*

<table>
<thead>
<tr>
<th>Stable Subscale</th>
<th>Internal Acute Subscale</th>
<th>External Acute Subscale</th>
<th>Protective Subscale</th>
</tr>
</thead>
<tbody>
<tr>
<td>Peer associations</td>
<td>Substance abuse</td>
<td>Interpersonal</td>
<td>Responsive to advice</td>
</tr>
<tr>
<td></td>
<td></td>
<td>relationships</td>
<td></td>
</tr>
<tr>
<td>Attitudes towards</td>
<td>Anger/hostility</td>
<td>Living situation</td>
<td>Prosocial identity</td>
</tr>
<tr>
<td>authority</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Impulse control</td>
<td>Negative mood</td>
<td>Employment</td>
<td>High expectations</td>
</tr>
<tr>
<td>Problem-solving</td>
<td></td>
<td>Attachment with others</td>
<td>Costs/benefits</td>
</tr>
<tr>
<td>Sense of entitlement</td>
<td></td>
<td></td>
<td>Social supports</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Opportunity/access</td>
<td></td>
<td></td>
<td>Social control</td>
</tr>
<tr>
<td>to victims</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Note.* The item *opportunity/access to victims* was an acute risk factor in the original DRAOR; the item *attachment with others* was originally a stable risk factor.
Figure 4

The relationship between treatment change and recidivism as mediated by Stable change

\[
a = .18^{**}
\]

\[
b = -.32^*
\]

\[
c' = -.02
\]
Endnotes

i Although, technically, without a measure of change over time for a control group, we cannot determine whether the change was brought about by the programme (Monahan & Skeem, 2014).

ii New Zealand legislation requires that all those released after two or more years in prison will be subject to post-release supervision (i.e., parole) of at least six months. While on parole, offenders must report regularly to their probation officer as well as adhering to a number of standard and/or special conditions (e.g., reside at an approved address, no contact with any victims or antisocial associates).

iii Based on the static risk instrument used by the New Zealand Department of Corrections: the RoC*RoI (Bakker, Riley, & O’Malley, 1999). The RoC*RoI is expressed as a probability and represents an offender’s estimated risk of reconviction leading to reimprisonment over five years in the community. RoC*RoI scores range from 0 (low) to 1.0 (very high).

iv In New Zealand, life-sentences can be imposed for murder, manslaughter or Class A drug dealing. Life-sentenced prisoners must serve a minimum period of 10 years, or 17 years in more severe cases, before they can be considered for parole; in the community they remain on parole for life and can be recalled to prison at any time.

v Breaches of parole and any reconviction were also examined but the VRS failed to predict these outcomes; for space considerations, they are not reported here.

vi Recall that “post-treatment” scores were updated for 31 men who remained in prison for more than 6 months after the programme.